

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:25:00 ; Search time 18.0843 Seconds
(without alignments)
582.095 Million cell updates/sec

Title: US-09-234-208B-1

Perfect score: 418
Sequence: 1 GTHSLPRPAVPLRMQP.....VGRGPPDAHVAVNLSREY 79

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	418	100.0	79	22	AAE09196 Human p68HER-2 ECD
2	418	100.0	83	22	AAE09185 Human p68HER-2 par
3	418	100.0	84	23	AAE20349 HER2 protein conta
4	413	98.8	79	22	AAE09188 Human p68HER-2 ECD
5	411	98.3	79	22	AAE09195 Human p68HER-2 ECD
6	411	98.3	79	22	AAE09199 Human p68HER-2 ECD
7	406	97.1	79	21	AAE09184 HER-2 C-terminal e
8	406	97.1	79	22	AAE09184 Human p68HER-2 ECD
9	406	97.1	79	23	AAE20350 Human HER2 intron
10	404	96.7	79	22	AAE09192 Human p68HER-2 ECD

11	403	96.4	79	22	AAE09190 Human p68HER-2 ECD
12	402	96.2	79	22	AAE09186 Human p68HER-2 ECD
13	400	95.7	79	22	AAE09189 Human p68HER-2 ECD
14	399	95.5	79	22	AAE09187 Human p68HER-2 ECD
15	399	95.5	79	22	AAE09191 Human p68HER-2 ECD
16	398	95.2	79	22	AAE09197 Human p68HER-2 ECD
17	398	95.2	79	22	AAE09198 Human p68HER-2 ECD
18	397	95.0	79	22	AAE09193 Human p68HER-2 ECD
19	396	94.7	79	22	AAE09194 Human p68HER-2 ECD
20	363	86.8	419	22	AAE09213 Human p68HER-2 gen
21	358	85.6	419	22	AAE09212 Human p68HER-2 gen
22	356	85.2	419	22	AAE09205 Human p68HER-2 gen
23	355	84.9	79	21	AAE09239 HER-2 C-terminal p
24	355	84.9	79	22	AAE09180 Human p68HER-2 ECD
25	355	84.9	79	23	AAE20347 Human p68HER-2 ECD
26	354	84.7	419	22	AAE09207 Human p68HER-2 gen
27	354	84.7	419	22	AAE09209 Human p68HER-2 gen
28	352	84.2	419	22	AAE09203 Human p68HER-2 gen
29	351	84.0	419	22	AAE09181 Human p68HER-2 gen
30	351	84.0	419	22	AAE09208 Human p68HER-2 gen
31	351	84.0	419	22	AAE09210 Human p68HER-2 gen
32	351	84.0	419	23	AAE20348 Human truncated HE
33	351	84.0	420	21	AAE09240 Human p68HER-2 gen
34	350	83.7	419	22	AAE09206 Human p68HER-2 gen
35	350	83.7	419	22	AAE09211 Human p68HER-2 gen
36	349	83.5	419	22	AAE09204 Human p68HER-2 gen
37	346	82.8	419	22	AAE09216 Human p68HER-2 gen
38	343	82.1	79	22	AAE09182 Human p68HER-2 ECD
39	343	82.1	419	22	AAE09200 Human p68HER-2 gen
40	343	82.1	419	22	AAE09202 Human p68HER-2 gen
41	339	81.1	419	22	AAE09183 Human p68HER-2 gen
42	337	80.6	419	22	AAE09214 Human p68HER-2 gen
43	337	80.6	419	22	AAE09215 Human p68HER-2 gen
44	86	20.6	459	20	AAE09215 HEV US and CMP-KDO
45	86	20.6	459	20	AAE09215 HEV US recombinant

ALIGNMENTS

RESULT 1
AAE09196
AAE09196 standard; peptide: 79 AA.
ID
AC AAE09196;
XX
DT 15-NOV-2001 (first entry)
XX
DE Human p68HER-2 ECDIIIA variant 11 encoded by HER-2 intron 8.
XX
KM HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;
KW p68HER-2; ECDIIIA; variant.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 6 /note="p68HER-2 ECDIIIA (AAE09184) Pro substituted
FT FT with Leu"
FT FT
FT Misc-difference 73 /note="p68HER-2 ECDIIIA (AAE09184) Asp substituted
FT FT with Asn which is encoded by CAC"
XX
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US05327.
XX
XX 16-FEB-2000; 2000US-0506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.

```

XX Clinton G, Henner WD, Evans A;
PI
XX
XX MPI: 2001-529934/58.
DR N-PSDB: AAD15865.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for
PI the treatment of hard tumors -
XX
XX Example 11: Page -: 61pp: English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 10-8. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is p68HER-2 ECDIIIA variant
CC Note: The present sequence is not shown in the specification but is
CC derived from HER-2 intron 8 encoded ECDIIIA sequence given in figure 8
CC (AAE09184).
XX
XX Sequence 79 AA:
SQ
XX
XX Query Match 100.0%; Score 418; DB 22; Length 79;
XX Best Local Similarity 100.0%; Pred. No. 2e-38;
XX Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 GTHSLPRPAANVPVLRMQPGPAHPVLSFLRPSWDLVSFAFYSLPLAPLSPTSPVPSV 60
OY |||||||
DB 1 GTHSLPRPAANVPVLRMQPGPAHPVLSFLRPSWDLVSFAFYSLPLAPLSPTSPVPSV 60
OY 61 GRGPPDAHAVNLSRYEG 79
DB 61 GRGPPDAHAVNLSRYEG 79
XX
XX RESULT 2
XX AAE09185 standard: Protein; 83 AA.
XX ID AAE09185;
XX AC AAE09185;
XX
XX 15-NOV-2001 (first entry)
XX
XX Human p68HER-2 partial protein containing ECDIIIA variant sequence.
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX solid tumour; cancer; polymorphism; cytostatic; gene therapy;
XX p68HER-2; ECDIIIA; variant.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH 3..81
XX FT Domain
XX FT /label= ECDIIIA
XX FT /note= "Extracellular domain IIA of p68HER-2"
XX FT Misc-difference 1
XX FT /note= "Encoded by CC"
XX FT Misc-difference 8
XX FT /note= "Most commonly occurring ECDIIIA (AAE09184) PRO
XX FT substituted with Leu"
XX FT Modified-site 75
XX FT /note= "Asn is N-glycosylated. Most commonly occurring
XX FT ECDIIIA (AAE09184) Asp substituted with Asn"
XX FT
XX

```

```

PN WO200161356-A1.
XX
XX 23-AUG-2001.
XX PD
XX 16-FEB-2001; 2001WO-US05327.
XX PF
XX 16-FEB-2000; 2000US-0506079.
XX PR
XX (UOOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
PI
XX
XX MPI: 2001-529934/58.
DR N-PSDB: AAD15854.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for
PI the treatment of hard tumors -
XX
XX Example 1: Fig 1: 61pp: English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 10-8. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is human p68HER-2 partial protein
CC containing ECDIIIA variant sequence.
XX
XX Sequence 83 AA:
SQ
XX
XX Query Match 100.0%; Score 418; DB 22; Length 83;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-38;
XX Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 GTHSLPRPAANVPVLRMQPGPAHPVLSFLRPSWDLVSFAFYSLPLAPLSPTSPVPSV 60
OY |||||||
DB 3 GTHSLPRPAANVPVLRMQPGPAHPVLSFLRPSWDLVSFAFYSLPLAPLSPTSPVPSV 62
OY 61 GRGPPDAHAVNLSRYEG 79
DB 63 GRGPPDAHAVNLSRYEG 81
XX
XX RESULT 3
XX AAE20349 standard: Protein; 84 AA.
XX ID AAE20349;
XX AC AAE20349;
XX
XX 18-JUN-2002 (first entry)
XX
XX HER2 protein containing extracellular domain (ECDIIIA).
XX
XX Endothelial growth factor receptor; EGFR; tumour; cytostatic; herstatin;
XX HER-2 receptor tyrosine kinase; squamous cell carcinoma; lung; colon;
XX glial cell tumour; cell growth.
XX
XX Unidentified.
XX
XX Key Location/Qualifiers
XX FH 1..2
XX FT Misc-difference 1.2
XX FT /note= "Encoded by CCGGA"
XX FT Modified-site 75..77
XX FT /note= "Asn is N-glycosylated"
XX FT Misc-difference 81..82
XX FT /note= "Encoded by GGCTGAGACGGCCCTTCCCCACCCACCCGCCAC
XX FT
XX

```

```

FT      TCCTCAGTC"
FT      Misc-difference 83..84
FT      /note="Encoded by TCGT"
PN      WO200214470-A2.
XX      21-FEB-2002.
XX      PD
XX      PF 14-AUG-2001; 2001WO-US25502.
XX      PR 14-AUG-2000; 2000US-0638834.
XX      PA (UIOR-) UNIV OREGON HEALTH SCI.
XX      PI Clinton GM.
XX      PI
XX      DR MPI: 2002-269185/31.
XX      DR N-PSDB: AAD32539.
XX      DR
XX      PT Treating solid tumor characterized by expression of endothelial growth
XX      PT factor receptor, involves administering recombinant herstatin that
XX      PT binds to extracellular domain of the endothelial growth factor receptor
XX      PT
XX      PS Example 11; Fig 1; 82pp; English.
XX      PS
XX      CC The present invention relates to a method for treating a solid tumour
XX      CC characterised by endothelial growth factor receptor (EGFR) expression.
XX      CC The method involves administering an agent that binds to an extracellular
XX      CC domain (ECD) of EGFR. The invention also relates to a naturally occurring
XX      CC inhibitor of HER-2 receptor tyrosine kinase called herstatin. The co-
XX      CC expression of herstatin with p185HER2 causes a striking reduction in cell
XX      CC growth that corresponds with suppression of p185 autophosphorylation. The
XX      CC method or a pharmaceutical composition is useful for treating a solid
XX      CC tumour (selected from squamous cell carcinoma, lung carcinoma, colon
XX      CC carcinoma and glioma cell tumour) characterised by EGFR expression. The
XX      CC present sequence is HER2 protein containing extracellular domain
XX      CC (ECDIIIA) sequence.
XX      CC
XX      SQ Sequence 84 AA:

Query Match 100.0%; Score 418; DB 23; Length 84;
Best Local Similarity 100.0%; Pred. No. 2.1e-38;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTHSLPPRAAVPVPLRMQPGPAHPVLSFLRPSMDVSAFVSLPLAPLSPTSPISPVSV 60
DB 3 GTHSLPPRAAVPVPLRMQPGPAHPVLSFLRPSMDVSAFVSLPLAPLSPTSPISPVSV 62
OY 61 GRGPPDAHVAVNLSRYEG 79
DB 63 GRGPPDAHVAVNLSRYEG 81

RESULT 4
AAE09188
ID AAE09188 standard; peptide; 79 AA.
AC AAE09188;
XX
XX 15-NOV-2001 (first entry)
XX
XX Human p68HER-2 ECDIIIA variant 3 encoded by HER-2 intron 8.
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX solid tumour; cancer; polymorphism; cytostatic; gene therapy;
XX p68HER-2; ECDIIIA; variant.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Misc-difference 6 /note="p68HER-2 ECDIIIA (AAE09184) Pro substituted
FT

```

```

FT      with Leu"
XX      WO200161356-A1.
XX      PN
XX      PD 23-AUG-2001.
XX      PF 16-FEB-2001; 2001WO-US05327.
XX      PR 16-FEB-2000; 2000US-0506079.
XX      PA (UIOR-) UNIV OREGON HEALTH SCI.
XX      PI Clinton G, Henner WD, Evans A;
XX      PI
XX      DR MPI: 2001-529934/58.
XX      DR N-PSDB: AAD15857.
XX      DR
XX      PT New polypeptide, which binds to the extracellular domain of HER-2 for
XX      PT the treatment of hard tumors -
XX      PT
XX      PS Example 11; Page -; 61pp; English.
XX      PS
XX      CC The invention relates to novel HER-2 (herstatin-2) antagonist
XX      CC particularly a polypeptide that binds to the extracellular domain (ECD)
XX      CC of HER-2 at a site that is different from the binding site of humanised
XX      CC antibody, Herceptin, at an affinity of at least 1078. The present
XX      CC invention is based upon the initial discovery of an alternative HER-2
XX      CC mRNA transcript with 274 bp insert of intron 8. The translation product
XX      CC of the alternative transcript is a truncated HER-2 protein designated
XX      CC p68HER-2 which lacks the transmembrane and intracellular domains of
XX      CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX      CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX      CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX      CC nucleic acids encoding these are useful to treat, diagnose and identify
XX      CC solid tumours. The present sequence is p68HER-2 ECDIIIA variant
XX      CC encoded by polymorphic form of human HER-2 intron 8.
XX      CC Note: The present sequence is not shown in the specification but is
XX      CC derived from HER-2 intron 8 encoded ECDIIIA sequence given in figure 8
XX      CC (AAE09184).
XX      CC
XX      SQ Sequence 79 AA:

Query Match 98.8%; Score 413; DB 22; Length 79;
Best Local Similarity 98.7%; Pred. No. 7e-38;
Matches 78; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTHSLPPRAAVPVPLRMQPGPAHPVLSFLRPSMDVSAFVSLPLAPLSPTSPISPVSV 60
DB 1 GTHSLPPRAAVPVPLRMQPGPAHPVLSFLRPSMDVSAFVSLPLAPLSPTSPISPVSV 60
OY 61 GRGPPDAHVAVNLSRYEG 79
DB 61 GRGPPDAHVAVNLSRYEG 79

RESULT 5
AAE09195
ID AAE09195 standard; peptide; 79 AA.
AC AAE09195;
XX
XX 15-NOV-2001 (first entry)
XX
XX Human p68HER-2 ECDIIIA variant 10 encoded by HER-2 intron 8.
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX solid tumour; cancer; polymorphism; cytostatic; gene therapy;
XX p68HER-2; ECDIIIA; variant.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Misc-difference 73
FT

```

FT		/note= "p68HER-2 ECDIIIA (AAE09184) Asp substituted
FT		with Asn which is encoded by CAC"
XX		
PN	WO200161356-A1.	
XX		
PD	23 -Aug-2001.	
XX		
PE	16-Feb-2001; 2001WO-US05327.	
XX		
PR	16-Feb-2000; 2000US-0506079.	
XX		
PA	(UYOR-) UNIV OREGON HEALTH SCI.	
XX		
PI	Clinton G, Henner WD, Evans A;	
XX		
DR	WPI: 2001-529934/58.	
DR	N-PSTDB: AAD15864.	
XX		
PT	New polypeptide, which binds to the extracellular domain of HER-2 for	
XX	the treatment of hard tumors -	
PS	Example 11: Page -: 61pp; English.	
XX		
CC	The invention relates to novel HER-2 (herstatin-2) antagonist	
CC	particularly a polypeptide that binds to the extracellular domain (ECD)	
CC	of HER-2 at a site that is different from the binding site of humanised	
CC	antibody, Herceptin, at an affinity of at least 10 ⁷ 8. The present	
CC	invention is based upon the initial discovery of an alternative HER-2	
CC	mRNA transcript with 274 bp insert of intron 8. The translation product	
CC	of the alternative transcript is a truncated HER-2 protein designated	
CC	p68HER-2 which lacks the transmembrane and intracellular domains of	
CC	p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.	
CC	The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise	
CC	the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the	
CC	nucleic acids encoding these are useful to treat, diagnose and identify	
CC	solid tumours. The present sequence is p68HER-2 ECDIIIA variant	
CC	encoded by polymorphic form of human HER-2 intron 8.	
CC	Note: The present sequence is not shown in the specification but is	
CC	derived from HER-2 intron 8 encoded ECDIIIA sequence given in figure 8	
CC	(AAE09184).	
XX		
SQ	Sequence 79 AA:	
	Query Match 98.3%; Score 411; DB 22; Length 79;	
	Best Local Similarity 98.7%; Pred. No. 1.2e-37;	
	Matches 78; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
OY	1 GTHTSLPRAAPVPLRMPGPAHPYLSPRLPSMDLVSATFSLPLAPLSPTSVIPSPVSV 60	
DB	1 GTHSLPRAAPVPLRMPGPAHPYLSPRLPSMDLVSATFSLPLAPLSPTSVIPSPVSV 60	
OY	61 GRGPDDAHVAVNLSRYEG 79	
DB	61 GRGPDDAHVAVNLSRYEG 79	
RESULT 6		
AAE09199		
ID	AAE09199 standard; peptide: 79 AA.	
XX		
AC	AAE09199;	
XX		
DT	15-NOV-2001 (first entry)	
XX		
DE	Human p68HER-2 ECDIIIA variant 15 encoded by HER-2 intron 8.	
XX		
KM	HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;	
KW	solid tumour; cancer; polymorphism; cytostatic; gene therapy;	
KM	p68HER-2; ECDIIIA; variant.	
XX		
OS	Homo sapiens.	
XX		
Key	Location/Qualifiers	
PH		

FT	Misc-difference	73	/note= "p68HER-2 ECDIIIA (AAE09184) Asp substituted with Asn"
FT			
XX			
XX			
XX			
XX			
PD	23-AUG-2001.		
XX			
XX			
PR	16-FEB-2001; 2001MO-USO5327.		
XX			
XX			
XX	16-FEB-2000; 2000US-0506079.		
PA	(UYOR-) UNIV OREGON HEALTH SCI.		
XX			
PI	Clinton G, Henner WD, Evans A;		
DR	WPI: 2001-529934/58.		
DR	N-PSDB: AAD15868.		
XX			
PT	New polypeptide, which binds to the extracellular domain of HER-2 for the treatment of hard tumors -		
XX			
PS	Example 12; Page -: 61pp; English.		
XX			
CC	The invention relates to novel HER-2 (herstatin-2) antagonist		
CC	particularly a polypeptide that binds to the extracellular domain (ECD)		
CC	of HER-2 at a site that is different from the binding site of humanised		
CC	antibody, Herceptin, at an affinity of at least 10 ⁷ 8. The present		
CC	invention is based upon the initial discovery of an alternative HER-2		
CC	mRNA transcript with 274 bp insert of intron 8. The translation product		
CC	of the alternative transcript is a truncated HER-2 protein designated		
CC	p68HER-2 which lacks the transmembrane and intracellular domains of		
CC	p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIA.		
CC	The ECDIIA-containing polypeptides bind tightly to, and thus antagonise		
CC	the HER-2 receptor. The peptides, which bind to an HFR-2 ECD, and the		
CC	nucleic acids encoding these are useful to treat, diagnose and identify		
CC	solid tumours. The present sequence is p68HER-2 ECDIIA variant		
CC	encoded by polymorphic form of human HER-2 intron 8.		
CC	Note: The present sequence is not shown in the specification but is		
CC	derived from HER-2 intron 8 encoded ECDIIA sequence given in figure 8		
CC	(AAE09184).		
XX			
SQ	Sequence	79 AA:	
Query Match	98.3%;	Score 411;	DB 22; Length 79;
Best Local Similarity	98.7%;	Pred. No. 1.2e-37;	
Matches	78; Conservative	0; Mismatches	1; Indels
			Gaps
OY	1	GHSLSLPPRAAPVPLBQPGPAHGVLSLRPSMWLVSAFYSLPLAPISPMSPVPSPVSV	60
Db	1	GHSLSLPPRAAPVPLBQPGPAHGVLSLRPSMWLVSAFYSLPLAPISPMSPVPSPVSV	60
OY	61	GRGPPDAHVAVNLSRYEG	79
Ddb	61	GRGPPDAHVAVNLSRYEG	79
RESULT 7			
ID	AA97241		
XX	AA97241 standard; Protein; 79 AA.		
AC	AA97241;		
DT	04-DEC-2000 (first entry)		
DE	HER-2 C-terminal extracellular domain IIIA.		
XX			
KM	HER-2; erbB-2; oncogene; receptor-like tyrosine kinase; insertion;		
KM	extracellular domain IIIA; antagonist; intron 8; C-terminal extension;		
KX	truncated HER-2; p68; dimerization inhibitor; cytostatic.		
OS	Homo sapiens.		
XX			


```
PN WO200044403-A1.
XX
PD 03-AUG-2000.
XX
PF 20-JAN-2000; 2000WO-US01484.
XX
PR 20-JAN-1999; 99US-0234208.
XX
PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
PI Doherty JK, Clinton GM, Adelman JP;
XX
DR WPI; 2000-499287/44.
XX
DR N-PSDB; AAA53783.
XX
PT Using polypeptides and antibodies that bind to the extracellular domain
PT of the receptor-like tyrosine kinase HER-2 to treat solid tumors of the
PT breast, lung, ovaries and colon
XX
PS Example 11; Page 42-43; 46pp; English.
XX
CC HER-2/neu (erbB-2) oncogene encodes a receptor-like tyrosine kinase. The
CC extracellular domain of p185-HER-2 is proteolytically shed from breast
CC carcinoma cells in culture and is found in serum of some cancer patients
CC and may be a serum marker of metastatic breast cancer. An alternative
CC HER-2 mRNA of 4.8 kb with a 274 bp insert (intron 8) has been
CC identified. The retained intron is in-frame and encodes a 79 amino acid
CC extension designated ECDIIIA (the present sequence), which is inserted at
CC residue 340 of p185-HER-2. The alternative mRNA predicts a truncated
CC HER-2 protein (approximately 66 kDa) that lacks the transmembrane and
CC intracellular domains (see AAY97240). p68HER-2 specifically binds to
CC p185-HER-2 without activating HER-2. It could therefore block
CC dimerization of p185-HER-2. The p68HER-2 polypeptide binds to a site on
CC the ECD of HER-2 that is different from the site of binding for
CC Hereceptin (RTM) (a marketed humanized monoclonal antibody that is used
CC for the treatment of cancer and binds to the ECD of HER-2). The methods,
CC compositions, polypeptides and antibodies are used to treat solid
CC tumours such as breast cancer, small cell lung carcinoma, ovarian cancer
CC and/or colon cancer, especially where over-expression of HER-2 is
CC indicated.
XX
SQ Sequence 79 AA:
XX
Query Match 97.1%; Score 406; DB 21; Length 79;
Best Local Similarity 97.5%; Pred. No. 4.1e-37;
Matches 77; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 GTHSLPRPAAYVPLRMQPGPAHPVLSFLRPSWDLVSFAFYSLLPLAPLSPTSPISPSV 60
DB 1 GTHSLPRPAAYVPLRMQPGPAHPVLSFLRPSWDLVSFAFYSLLPLAPLSPTSPISPSV 60
QY 61 GRGPDPDAHVAVNLRYEG 79
DB 61 GRGPDPDAHVAVNLRYEG 79
XX
RESULT 8
AAE09184
ID AAE09184 standard; peptide; 79 AA.
XX
AC AAE09184;
XX
DT 15-NOV-2001 (first entry)
XX
DE Human p68HER-2 ECDIIIA domain encoded by HER-2 intron 8.
XX
KW HER-2; herstatin; antagonist; extracellular domain; ECD; Hereceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;
KW p68HER-2; ECDIIIA; intron 8.
XX
OS Homo sapiens.
XX
PR WO200161356-A1.
PN
```

```
XX
PD 23-AUG-2001.
XX
PF 16-FEB-2001; 2001WO-US05327.
XX
PR 16-FEB-2000; 2000US-0506079.
XX
PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
PI Clinton G, Henner WD, Evans A;
XX
DR WPI; 2001-529934/58.
XX
DR N-PSDB; AAD15853, AAD15869.
XX
PT New polypeptide, which binds to the extracellular domain of HER-2 for
PT the treatment of hard tumors -
XX
PS Example 11; Fig 8; 61pp; English.
XX
CC The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Hereceptin, at an affinity of at least 10-8. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is ECDIIIA domain of human p68HER-2
CC protein encoded by intron 8 of HER-2 gene.
XX
SQ Sequence 79 AA:
XX
Query Match 97.1%; Score 406; DB 22; Length 79;
Best Local Similarity 97.5%; Pred. No. 4.1e-37;
Matches 77; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 GTHSLPRPAAYVPLRMQPGPAHPVLSFLRPSWDLVSFAFYSLLPLAPLSPTSPISPSV 60
DB 1 GTHSLPRPAAYVPLRMQPGPAHPVLSFLRPSWDLVSFAFYSLLPLAPLSPTSPISPSV 60
QY 61 GRGPDPDAHVAVNLRYEG 79
DB 61 GRGPDPDAHVAVNLRYEG 79
XX
RESULT 9
AAE20350
ID AAE20350 standard; Protein; 79 AA.
XX
AC AAE20350;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human HER2 intron 8 encoded protein.
XX
KW Human; tumour; endothelial growth factor receptor; EGFR; cytostatic;
KW herstatin; HER-2 receptor tyrosine kinase; squamous cell carcinoma; lung;
KW colon; glial cell tumour; cell growth.
XX
OS Homo sapiens.
XX
PN WO200214470-A2.
XX
PD 21-FEB-2002.
XX
PF 14-AUG-2001; 2001WO-US25502.
XX
PR 14-AUG-2000; 2000US-0638834.
PN
```

PA (UYOR-) UNIV OREGON HEALTH SCI.
 XX
 PI Clinton GW.
 XX
 DR WPI: 2002-269185/31.
 DR N-PSDB: AAD32540.

XX Treating solid tumor characterized by expression of endothelial growth
 PT factor receptor, involves administering recombinant herstatin that
 PT binds to extracellular domain of the endothelial growth factor receptor

PT Example 11: Fig 8: 82pp: English.

XX The present invention relates to a method for treating a solid tumour
 CC characterised by endothelial growth factor receptor (EGFR) expression.
 CC The method involves administering an agent that binds to an extracellular
 CC domain (ECD) of EGFR. The invention also relates to a naturally occurring
 CC inhibitor of HER-2 receptor tyrosine kinase called herstatin. The co-
 CC expression of herstatin with p185HER2 causes a striking reduction in cell
 CC growth that corresponds with suppression of p185 autophosphorylation. The
 CC method or a pharmaceutical composition is useful for treating a solid
 CC tumour (selected from squamous cell carcinoma, lung carcinoma, colon
 CC carcinoma and glial cell tumour) characterised by EGFR expression. The
 CC present sequence is a protein encoded by human HER2 intron 8.

XX Sequence 79 AA:

Query Match 97.1%; Score 406; DB 23; Length 79;
 Best Local Similarity 97.5%; Pred. No. 4.1e-37;
 Matches 77; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GTHSLPRRAAVVPLRMOPGPAHPVLSFLRPSMDLSAFYSIPLAPLPTSPVISPVS 60
 DB 1 GTHSLPRRAAVVPLRMOPGPAHPVLSFLRPSMDLSAFYSIPLAPLPTSPVISPVS 60
 OY 61 GRGPPDAHVAANLSRYEG 79
 DB 61 GRGPPDAHVAANLSRYEG 79

RESULT 10
 AAE09192
 ID AAE09192 standard; peptide: 79 AA.

XX AAE09192;

DT 15-NOV-2001 (first entry)

DE Human p68HER-2 ECDIIIA variant 7 encoded by HER-2 intron 8.

XX HER-2; herstatin; antagonist; extracellular domain; ECD; Hereceptin;
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;
 KW p68HER-2; ECDIIIA; variant.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 36 /note="p68HER-2 ECDIIIA (AAE09184) Leu substituted
 FT with Ile"

PN WO200161356-A1.

PD 23-AUG-2001.

PF 16-FEB-2001; 2001WO-US05327.

PR 16-FEB-2000; 2000US-0506079.

PA (UYOR-) UNIV OREGON HEALTH SCI.

XX Clinton G, Henner WD, Evans A;

XX WPI: 2001-529934/58.
 DR N-PSDB: AAD15861.

XX New polypeptide, which binds to the extracellular domain of HER-2 for
 PT the treatment of hard tumors -

XX Example 11: Page -: 61pp: English.

XX The invention relates to novel HER-2 (herstatin-2) antagonist
 CC particularly a polypeptide that binds to the extracellular domain (ECD)
 CC of HER-2 at a site that is different from the binding site of humanised
 CC antibody, Herceptin, at an affinity of at least 10⁸. The present
 CC invention is based upon the initial discovery of an alternative HER-2
 CC mRNA transcript with 274 bp insert of intron 8. The translation product
 CC of the alternative transcript is a truncated HER-2 protein designated
 CC p68HER-2 which lacks the transmembrane and intracellular domains of
 CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
 CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
 CC nucleic acids encoding these are useful to treat, diagnose and identify
 CC solid tumours. The present sequence is p68HER-2 ECDIIIA variant
 CC encoded by polymorphic form of human HER-2 intron 8.
 CC Note: The present sequence is not shown in the specification but is
 CC derived from HER-2 intron 8 encoded ECDIIIA sequence given in figure 8
 CC (AAE09184).

XX Sequence 79 AA:

Query Match 96.7%; Score 404; DB 22; Length 79;
 Best Local Similarity 96.2%; Pred. No. 6.8e-37;
 Matches 76; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GTHSLPRRAAVVPLRMOPGPAHPVLSFLRPSMDLSAFYSIPLAPLPTSPVISPVS 60
 DB 1 GTHSLPRRAAVVPLRMOPGPAHPVLSFLRPSMDLSAFYSIPLAPLPTSPVISPVS 60
 OY 61 GRGPPDAHVAANLSRYEG 79
 DB 61 GRGPPDAHVAANLSRYEG 79

RESULT 11
 AAE09190
 ID AAE09190 standard; peptide: 79 AA.

XX AAE09190;

DT 15-NOV-2001 (first entry)

DE Human p68HER-2 ECDIIIA variant 5 encoded by HER-2 intron 8.

XX HER-2; herstatin; antagonist; extracellular domain; ECD; Hereceptin;
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;
 KW p68HER-2; ECDIIIA; variant.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 18 /note="p68HER-2 ECDIIIA (AAE09184) Met substituted
 FT with Leu which is encoded by A7A"

PN WO200161356-A1.

PD 23-AUG-2001.

PF 16-FEB-2001; 2001WO-US05327.

PR 16-FEB-2000; 2000US-0506079.

PA (UYOR-) UNIV OREGON HEALTH SCI.

XX

```
PI Clinton G, Henner WD, Evans A;
XX
XX WPI: 2001-529934/58.
DR N-PSDB; AAD15859.
XX
PT New polypeptide, which binds to the extracellular domain of HER-2 for
PT the treatment of hard tumors -
XX
PS Example 11: Page -: 61pp: English.
XX
CC The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 10-8. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD 1, II of the p185HER-2 and the novel ECDIIIIa.
CC The ECDIIIIa-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is p68HER-2 ECDIIIIa variant
CC encoded by polymorphic form of human HER-2 intron 8.
CC Note: The present sequence is not shown in the specification but is
CC derived from HER-2 intron 8 encoded ECDIIIIa sequence given in figure 8
CC (AAE09184).
XX
SQ Sequence 79 AA:
XX
Query Match 96.4%; Score 403; DB 22; Length 79;
Best Local Similarity 96.2%; Pred. No. 8.7e-37;
Matches 76; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 1 GTHSLPPRAAAYVPLRMQGPAPHPVLSFLRPSWDLVSAFYSLPLAPLSPISPVSV 60
DB 1 GTHSLPPRAAAYVPLRLDPGAPHPVLSFLRPSWDLVSAFYSLPLAPLSPISPVSV 60
OY 61 GRGPPDAHVAANLSRYEG 79
DB 61 GRGPPDAHVAANVLSRYEG 79
RESULT 12
AAE09186
ID AAE09186 standard; peptide: 79 AA.
XX
AC AAE09186;
XX
DT 15-NOV-2001 (first entry)
XX
DE Human p68HER-2 ECDIIIIa variant 1 encoded by HER-2 intron 8.
XX
KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;
KW p68HER-2; ECDIIIIa; variant.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT Misc-difference 2 /note="p68HER-2 ECDIIIIa (AAE09184) Thr substituted
XX FT with Ser."
XX PN WO200161356-A1.
XX PD 23-AUG-2001.
XX PF 16-FEB-2001; 2001WO-US05327.
XX PR 16-FEB-2000; 2000US-0506079.
XX PA (UYOR-) UNIV OREGON HEALTH SCI.
```

```
XX
XX Clinton G, Henner WD, Evans A;
XX
XX WPI: 2001-529934/58.
DR N-PSDB; AAD15855.
XX
PT New polypeptide, which binds to the extracellular domain of HER-2 for
PT the treatment of hard tumors -
XX
PS Example 11: Page -: 61pp: English.
XX
CC The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 10-8. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIIa.
CC The ECDIIIIa-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is p68HER-2 ECDIIIIa variant
CC encoded by polymorphic form of human HER-2 intron 8.
CC Note: The present sequence is not shown in the specification but is
CC derived from HER-2 intron 8 encoded ECDIIIIa sequence given in figure 8
CC (AAE09184).
XX
SQ Sequence 79 AA:
XX
Query Match 96.2%; Score 402; DB 22; Length 79;
Best Local Similarity 96.2%; Pred. No. 1.1e-36;
Matches 76; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 1 GTHSLPPRAAAYVPLRMQGPAPHPVLSFLRPSWDLVSAFYSLPLAPLSPISPVSV 60
DB 1 GTHSLPPRAAAYVPLRLDPGAPHPVLSFLRPSWDLVSAFYSLPLAPLSPISPVSV 60
OY 61 GRGPPDAHVAANLSRYEG 79
DB 61 GRGPPDAHVAANVLSRYEG 79
RESULT 13
AAE09189
ID AAE09189 standard; peptide: 79 AA.
XX
AC AAE09189;
XX
DT 15-NOV-2001 (first entry)
XX
DE Human p68HER-2 ECDIIIIa variant 4 encoded by HER-2 intron 8.
XX
KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;
KW p68HER-2; ECDIIIIa; variant.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT Misc-difference 16 /note="p68HER-2 ECDIIIIa (AAE09184) Leu substituted
XX FT with Gln"
XX PN WO200161356-A1.
XX PD 23-AUG-2001.
XX PF 16-FEB-2001; 2001WO-US05327.
XX PR 16-FEB-2000; 2000US-0506079..
XX PA
```

```
PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX
PI Clinton G, Henner WD, Evans A;
XX
DR WPI: 2001-529934/58.
XX N-PSDB: AAD15858.
XX
PT New polypeptide, which binds to the extracellular domain of HER-2 for
XX the treatment of hard tumors -
XX
PS Example 11: Page -: 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Hereceptin, at an affinity of at least 10-8. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is p68HER-2 ECDIIIA variant
XX encoded by polymorphic form of human HER-2 intron 8.
XX Note: The present sequence is not shown in the specification but is
XX derived from HER-2 intron 8 encoded ECDIIIA sequence given in figure 8
XX (AAE09184).
XX
SQ Sequence 79 AA:
XX
XX Query Match 95.7%; Score 400; DB 22; Length 79;
XX Best Local Similarity 96.2%; Pred. No. 1.9e-36;
XX Matches 76; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 GTHSLPRRAAVVPLRMOPGPAHPVLSFLRPSMDLVSAFYSILPLAPLSPTSPISPSV 60
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 GTHSLPRRAAVVPLRMOPGPAHPVLSFLRPSMDLVSAFYSILPLAPLSPTSPISPSV 60
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GRGPPDAHVAVNLSTRYEG 79
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 GRGPPDAHVAVNLSTRYEG 79
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
AAE09187
ID AAE09187 standard; peptide; 79 AA.
XX
XX AAE09187;
XX
AC 15-NOV-2001 (first entry)
XX
DT Human p68HER-2 ECDIIIA variant 2 encoded by HER-2 intron 8.
XX
DE HER-2; herstatin; antagonist; extracellular domain; ECD; Hereceptin;
XX solid tumour; cancer; polymorphism; cytostatic; gene therapy;
XX p68HER-2; ECDIIIA; variant.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT Misc-difference 5 /note="p68HER-2 ECDIIIA (AAE09184) Leu substituted
XX FT with Pro"
XX
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US05327.
XX
XX 16-FEB-2000; 2000US-0506079.
XX
```

```
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX
PI Clinton G, Henner WD, Evans A;
XX
DR WPI: 2001-529934/58.
XX N-PSDB: AAD15856.
XX
PT New polypeptide, which binds to the extracellular domain of HER-2 for
XX the treatment of hard tumors -
XX
PS Example 11: Page -: 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Hereceptin, at an affinity of at least 10-8. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is p68HER-2 ECDIIIA variant
XX encoded by polymorphic form of human HER-2 intron 8.
XX Note: The present sequence is not shown in the specification but is
XX derived from HER-2 intron 8 encoded ECDIIIA sequence given in figure 8
XX (AAE09184).
XX
SQ Sequence 79 AA:
XX
XX Query Match 95.5%; Score 399; DB 22; Length 79;
XX Best Local Similarity 96.2%; Pred. No. 2.4e-36;
XX Matches 76; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 GTHSLPRRAAVVPLRMOPGPAHPVLSFLRPSMDLVSAFYSILPLAPLSPTSPISPSV 60
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 GTHSLPRRAAVVPLRMOPGPAHPVLSFLRPSMDLVSAFYSILPLAPLSPTSPISPSV 60
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GRGPPDAHVAVNLSTRYEG 79
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 GRGPPDAHVAVNLSTRYEG 79
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
AAE09191
ID AAE09191 standard; peptide; 79 AA.
XX
XX AAE09191;
XX
AC 15-NOV-2001 (first entry)
XX
DT Human p68HER-2 ECDIIIA variant 6 encoded by HER-2 intron 8.
XX
DE HER-2; herstatin; antagonist; extracellular domain; ECD; Hereceptin;
XX solid tumour; cancer; polymorphism; cytostatic; gene therapy;
XX p68HER-2; ECDIIIA; variant.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT Misc-difference 21 /note="p68HER-2 ECDIIIA (AAE09184) Gly substituted
XX FT with Asp, Ala or Val"
XX
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US05327.
XX
XX
```

PR 16-FEB-2000; 2000US-0506079.

XX
PA (UYOR-) UNIV OREGON HEALTH SCI.
UN

PI Clinton G, Henner WD, Evans A;

DR WPI; 2001-529934/58.
DR N-PSDB: AAD15860

DR N-PSDB; AAD15860.

PT New polypeptide, which binds to the extracellular domain of HER-2 for
PT the treatment of hard tumors -

PS Example 11; Page -; 61pp; English.

CC The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 10⁸. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p66HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD 1. II of the p185HER-2 and the novel ECDIIIA.
CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is p66HER-2 ECDIIIA variant
CC encoded by polymorphic form of human HER-2 intron 8.
CC Note: The present sequence is not shown in the specification but is
CC derived from HER-2 intron 8 encoded ECDIIIA sequence given in Figure 8
CC (AAE09184).

50 Sequence 79 AA;

Query Match	95.58;	Score 399;	DB 22;	Length 79;
Best Local Similarity	06.39;	Prod No 3	48-36	

Best Local Similarity 96.28; Pred. No. 2.4e-36;

Matches	76;	Conservative	1;	Mismatches	2;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

OY 1 GTHSLPRPAVPVPLRMQPGRAHPVLSELRSDWLSAFYSLPLAPLSPTSVPSPSV 600
|||||
Db 1 GTHTLP RPRAVPVPLRMQGPXPAHPVLSELRSDWLSAFYSLPLAPLSPTSVPSPSV 600

```

OY      61 GRGPPDAHVAVNLSRYEG 79
          |||||:|||||
Db      61 GRGPPDAHVAVDLSRYEG 79

```

```
Search completed: March 4, 2003, 12:34:31
Job time : 19.0843 secs
```



```

A:Residues: 1-834 <REV>
A:Cross-references: EMBL:AL0031986
A:Experimental source: cultivar Columbia; BAC clone FAB14
R:Nawrath, C.; Schell, J.; Koncz, C.
Mol. Gen. Genet. 223, 65-75, 1990
A>Title: Homologous domains of the largest subunit of eucaryotic RNA polymerase II ar
A:Reference number: SI2071; MUID:91080867; PMID:2259344
A:Accession: SI2071
A:Molecule type: DNA
A:Residues: 1-421,'S','423-732','D','734-1055','R','1057-1714','SPYSPSY','1715-1834 <NAM1>
A:Cross-references: EMBL:X52954; NID:916504; PIDN:CMA37130.1; PID:916505
A:Experimental source: cv. Columbia
A>Note: the authors translated the codon AGC for residue 1755 as Arg
A:Accession: S27346
A:Molecule type: mRNA
A:Residues: 510-732,'D','734-1055','R','1057-1714','SPYSPSY','1715-1834 <NAM2>
A:Experimental source: cv. Columbia
R:Dietrich, M.A.; Prenger, J.P.; Guilfoyle, T.J.
Plant Mol. Biol. 15, 207-223, 1990
A>Title: Analysis of the genes encoding the largest subunit of RNA polymerase II in A
A:Reference number: SI1960; MUID:91355869; PMID:2103447
A:Accession: SI1960
A:Molecule type: DNA
A:Residues: 1-116,125-192,'NSKEE','198-297','R','299-302','R','304-400','KE','403','VDYGPHPPP
A:Cross-references: EMBL:X52494; NID:916493; PIDN:CMA36735.1; PID:916494
A>Note: the authors translated the codon CCT for residue 1083 as Ala
C:Genetics:
A:Gene: rpl1215; RPB1
A:Map position: 4
A:Introns: 28/3; 123/3; 218/3; 271/3; 325/3; 411/1; 440/3; 552/3; 648/2; 734/3; 1755/5
A:Note: FAB14.70
C:Superfamily: human DNA-directed RNA polymerase II largest chain
C:Keywords: DNA binding; nucleotidyltransferase; phosphoprotein; tandem repeat; trans
F:66-109/Region: zinc finger CCCC motif
F:1531-1812/Region: 7-residue repeats

Query Match          19.1%: Score 80; DB 1; Length 1834;
Best Local Similarity 36.4%: Pred. No. 7.6;
Matches 28; Conservative 3; Mismatches 36; Indels 10; Gaps 3;

Oy      7 PPAAPVPLRLNQP----GPAHPVLFLRPSPMDVLSAFSLPLAPLPTSPVSVGYGR 62
        | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     1730 PSIAIAPSNAIRISPSAPSPSPNSPSPSPSPSPSPSPSPSPSPSPSPSPYS-----PSSPTFSPPSSPYSSGA 1785

Oy      63 GPD--PDAHVAVNLSTRY 77
        ||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db     1786 SPDYSPSAGYSPTLPGLY 1802


RESULT 5
GB5422
hypothetical protein AT4g35800 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001
C:Accession: GB5422
R:anonymus, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp
Nature 402, 769-777, 1999
A>Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: AB5001; MUID:20083488; PMID:10617198
A:Accession: GB5422
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1840 <STO>
A:Cross-references: GB:NC_001268; NID:g7270532; PIDN:CAB81489.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4g35800
A:Map position: 4
C:Superfamily: human DNA-directed RNA polymerase II largest chain

Query Match          19.1%: Score 80; DB 2; Length 1840;
Best Local Similarity 36.4%: Pred. No. 7.7;
Matches 28; Conservative 3; Mismatches 36; Indels 10; Gaps 3;

```


ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: A96826

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-356 <STO>

A:Cross-references: GB:AE005173; NID:g4835761; PIDN:AA030228.1; GSPDB:GN00141

C:Genetics:

A:Gene: T8K14.10

A:Map position: 1

Query Match

Best Local Similarity 18.2%; Score 76; DB 2; Length 356;
Pred. No. 2.9; Mismatches 30; Indels 12; Gaps 3;

Matches 26; Conservative 7; Mismatches 30; Indels 12; Gaps 3;

QY 9 PAAPVPLRMQPGPAHVPVLSFLRPSMDVSAF-----YSLPLAPLSPTSPISV---VSV 60

DB 216 PGVPLGPPYSEPGSPPTGSPSPS---SGFLPPIYPPPMAPSPSVTPTSAYWCVAK 271

QY 61 GRGPPDAHVAVNLS 75

DB 272 PSYPPDPIQEAAMFA 286

RESULT 11

S48273

Probable transcription factor YBR108w - Yeast (*Saccharomyces cerevisiae*)

N:Alternate names: hypothetical protein YBR0901

C:Species: *Saccharomyces cerevisiae*

C>Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 19-Apr-2002

C:Accession: S48273; S45976; S44688

R:Manhaupt, G.; Stucka, R.; Ehme, S.; Vetter, I.; Feldmann, H.

Yeast 10.1363-1381.1994

A:Title: Analysis of a 70 kb region on the right arm of yeast chromosome II.

A:Reference number: S48255; MUID:95208357; PMID:7900426

A:Accession: S48273

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-848 <MAN>

A:Cross-references: EMBL:X78993; NID:g476045; PIDN:CA55611.1; PID:g476064

R:Feldmann, H.; Manhaupt, G.; Schwarzlose, C.; Vetter, I.

submitted to the Protein Sequence Database, August 1994

A:Reference number: S45927

A:Accession: S45976

A:Molecule type: DNA

A:Residues: 1-848 <FE2>

A:Cross-references: EMBL:235977; NID:g536378; PID:g536379; MIPS:YBR108w

C:Genetics:

A:Cross-references: SCD:S0000312

A:Map position: 2R

Query Match

Best Local Similarity 18.2%; Score 76; DB 2; Length 848;
Pred. No. 7.9; Mismatches 28; Indels 10; Gaps 3;

Matches 23; Conservative 6; Mismatches 28; Indels 10; Gaps 3;

QY 13 PVLRLMQPGPAHVPV---LSFLRPSMDVSAF---FVSLPLAPLSPTSPV---ISPSVGR 62

DB 383 PVLVRMQPGPGPGMGQGNITPIEPSLSDSTGTFHEVTPDPDAAPKPIDIPYDVSS 442

QY 63 GPPDDAH 69

DB 443 LPBPPTN 449

RESULT 12

D49600

genome-linked protein VPg - soybean dwarf virus

C:Species: soybean dwarf virus

C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Sep-1999

C:Accession: D49600

R:Ratjen, J.P.; Karageorgios, L.E.; Habili, N.; Waterhouse, P.M.; Symons, R.H.

Virology 198, 671-679, 1994

A:Title: Soybean dwarf luteovirus contains the third variant genome type in the luteo

A:Reference number: A49600; MUID:94120742; PMID:8291248

A:Accession: D49600

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-189 <RAT>

A:Cross-references: GB:L24049; NID:g436017; PIDN:AA17538.1; PID:g436021

C:Superfamily: potato leaf roll virus genome-linked protein

Query Match

Best Local Similarity 17.9%; Score 75; DB 2; Length 189;
Pred. No. 1.8; Mismatches 22; Indels 8; Gaps 2;

Matches 19; Conservative 12; Mismatches 22; Indels 8; Gaps 2;

QY 2 THSLRPAAVAVPVL-----RMQPGPAHVPVLSFLRPSMDVSAF---AFYSLPLAPLSPTSV 53

DB 58 THSCPGRTASMYPREVLSGRLYONASLSLMEYRSPTMINISRYSSSSRPLPPROY 117

QY 54 P 54

DB 118 P 118

RESULT 13

T48814

hypothetical protein 15E6.220 [Imported] - *Neurospora crassa*

C:Species: *Neurospora crassa*

C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 05-May-2000

C:Accession: T48814

R:Schulte, U.; Algn, V.; Hohensel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu

submitted to the Protein Sequence Database, April 2000

A:Reference number: 224541

A:Accession: T48814

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1952 <SCH>

A:Cross-references: EMBL:AL353822; GSPDB:GN00112; NCSP:15E6.220

A:Experimental source: cosmid contig 15E6; strain 74

C:Genetics:

A:Gene: NCSP:15E6.220

A:Map position: 2

A:introns: 281/3

Query Match

Best Local Similarity 17.9%; Score 75; DB 2; Length 1952;
Pred. No. 26; Mismatches 31; Indels 6; Gaps 3;

Matches 24; Conservative 8; Mismatches 31; Indels 6; Gaps 3;

QY 2 THSLRPAAVAVPVLRLMQPGPAHVPVLSFLRPSMDVSAF---ISPLAPLSPTSPV---ISPV 58

DB 1781 TPQAVPRPSTSLTPRAOPGPVSPAVS--GSGVAPASQAQSVAPAPVSSSTPVPAATVAPA 1838

QY 59 S-VGRGPPD 66

DB 1839 STVAAAPTP 1847

RESULT 14

T11029

beta-adaptin homolog F8L21.170 - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 24-Nov-1999

C:Accession: T11029

R:Bevan, M.; Peters, S.A.; van Staveren, M.; Dirse, W.; Stiekema, W.; Bancroft, I.

submitted to the Protein Sequence Database, July 1999

A:Reference number: 217587

A:Accession: T11029

A:Molecule type: DNA

A:Residues: 1-894 <BEV>

A:Cross-references: EMBL:AL096882; GSPDB:GN00062; ATSP:F8L21.170

A:Experimental source: cultivar Columbia; BAC clone F8L21

C:Genetics:

A:Gene: ATSP:F8L21.170

A:Map position: 4

A:introns: 29/3; 95/3; 133/3; 217/2; 262/3; 322/3; 353/3; 424/2; 445/3; 505/3; 536/3;

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:25:15 ; Search time 2.85542 Seconds

(without alignments)
1147.512 Million cell updates/sec

Title: US-09-234-208b-1

Perfect score: 418
Sequence: 1 GTHSLLPRAAVPVLRMQP.....YGRGPDPAHVAVNLRYEG 79

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83.5	20.0	888	SM6B_HUMAN	Q9h1t3 homo sapien
2	81.5	19.5	1257	PGCN_RAT	P53067 rattus norv
3	80	19.1	428	FXB2_MOUSE	Q64733 mus musculu
4	80	19.1	1840	RPB1_ARATH	P18616 arabidopsi
5	78	18.7	2319	NTC3_RAT	Q91172 rattus norv
6	76	18.2	848	YBVB_YEAST	P38266 saccharomyc
7	74.5	17.8	1255	PER2_HUMAN	O15055 homo sapien
8	73.5	17.6	864	YCI8_HUMAN	Q9u1k2 homo sapien
9	73.5	17.6	1822	SM6B_RAT	P49750 homo sapien
10	73.5	17.2	440	NK4R_HUMAN	P30098 homo sapien
11	72	17.2	828	SOX6_HUMAN	P35712 homo sapien
12	71.5	17.1	2441	CBP_MOUSE	P43481 mus musculu
13	71.5	16.9	291	YD53_SYNY3	P42350 synchocyst
14	70.5	16.9	788	DPOL_HPBHE	P13346 heron hepat
15	70.5	16.9	1268	PGCN_MOUSE	P55066 mus musculu
16	70.5	16.7	2318	NTC3_MOUSE	Q61982 mus musculu
17	70	16.7	2321	NTC3_HUMAN	Q9un47 homo sapien
18	69.5	16.6	433	TEA3_CHICK	O9u071 gallus gall
19	69.5	16.5	265	CDX1_HUMAN	P54951 mus musculu
20	68.5	16.4	886	SM6B_MOUSE	O15234 homo sapien
21	68.5	16.3	703	ML51_HUMAN	O05859 mus musculu
22	68	16.3	1206	FM14_MOUSE	O05860 mus musculu
23	68	16.3	1468	FMN1_MOUSE	Q05860 mus musculu
24	68	16.3	1752	RPB1_SCHPO	P36594 schizosacch
25	68	16.3	1852	RPB1_CAEEL	P16356 caenorhabd
26	68	16.3	3298	PC16_HUMAN	Q96190 homo sapien
27	67.5	16.1	123	VST1_HEVBU	P29325 hepatitis e
28	67.5	16.1	123	VST1_HEVBU	O04612 hepatitis e
29	67.5	16.1	123	VST1_HEVBU	P02817 bos taurus
30	67.5	16.1	213	AMEX_BOVIN	P50428 mus musculu
31	67.5	16.1	506	ARSA_MOUSE	P70178 mus musculu
32	67.5	16.1	667	SIX5_MOUSE	P40645 mus musculu
33	67.5	16.1	827	SOX6_MOUSE	

34	67	16.0	360	1	A2HS_RABIT	P80191 oryctolagus
35	67	16.0	817	1	VRP1_YEAST	P37370 saccharomyc
36	67	16.0	1433	1	Y310_HUMAN	O15027 homo sapien
37	66.5	15.9	259	1	MSP8_EIMAC	P09125 eimeria ace
38	66.5	15.9	283	1	EXTN_SORBI	P24152 sorghum bic
39	66.5	15.9	322	1	RX1_XENLA	O42201 xenopus lae
40	66.5	15.9	435	1	TEA3_HUMAN	O99594 homo sapien
41	66.5	15.9	1229	1	NI21_HUMAN	O9Y2n3 homo sapien
42	66	15.8	558	1	TF65_CHICK	P98152 gallus gall
43	66	15.8	736	1	ORP9_HUMAN	O965u4 homo sapien
44	66	15.8	868	1	NRG2_RAT	O35569 rattus norv
45	66	15.8	1447	1	DCC_HUMAN	P43146 homo sapien

ALIGNMENTS

RESULT 1
ID SM6B_HUMAN STANDARD: PRT; 888 AA.
AC Q9H3T3; Q9NRK9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Semaphorin 6B precursor (Semaphorin Z) (Sema Z).
GN SEMA6B OR SEMAZ.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RA Kimura T., Ishida H.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE-21248680; PubMed-11350127;
RA Correa R.G., Sasahara R.M., Bengtson M.H., Katayama M.L.H.,
RA Salim A.C.M., Brentani M.M., Sogayar M.C., de Souza S.J.,
RA Simpson A.J.G.;
Human semaphorin 6b.;
Genomics 73:343-348(2001).
CC -!- FUNCTION: MAY PLAY A ROLE IN BOTH PERIPHERAL AND CENTRAL NERVOUS
CC -!- SYSTEM DEVELOPMENT (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS: At least 2 isoforms; 1 (shown here) and
CC 2/6b.1; are produced by alternative splicing.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB022433; BAB20669.1; -
CC EMBL; AF216389; AAF87661.1; -
CC Genew: HGNC:10739; SEMA6B.
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR001627; Sema.
DR Pfam: PF01403; Sema; 1.
DR SMART: SM00423; PSI; 1.
DR SMART: SM00286; PTI; 1.
KW Signal. Transmembrane. Multigene family. Neurogenesis; Glycoprotein;
KW Developmental protein; Alternative splicing.
FT SIGNAL 1 25
FT CHAIN 26 888 SEMAPHORIN 6B.
FT DOMAIN 26 603 EXTRACELLULAR (POTENTIAL).

```

FT TRANSMEM 604 624 POTENTIAL.
FT DOMAIN 625 888 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 238 547 SEMA.
FT DOMAIN 661 674 POLY-GLY.
FT DOMAIN 750 753 POLY-LEU.
FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 155 155 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 167 167 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 462 462 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPPLIC 471 517 LSVFLFEETFRPDRCGRGGEGTCGLLSLEDAASGGL
AAFPRC -> RVCVGHACACVCHERSRSPQRPGRWLSRR
MGFORARGPORCRVY (IN ISOPFORM 1).
FT FT MISSING (IN ISOPFORM 2).
FT VARSPPLIC 518 888 D -> E (IN REF. 2).
FT CONFLICT 30 30
SQ SEQUENCE 888 AA: 95270 MM: 6FFB4D6828C70CB CRC64:

Query Match 20.0%; Score 83.5; DB 1; Length 888;
Best Local Similarity 35.4%; Pred. No. 0.8;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;

OY 1 GTH-----SLRPAAVVPRLMPGPAHPLSLRP-SMD---LVSAFYSLPLAPLSP 51
DB 698 GPHDLDSGLLPTEQHPRLPQRLPRLP-HPHALGPRAWDHGHPLPASASSLLAPRA 756
OY 52 SVIPSPVSGRGPDPDAH 70
DB 757 RAPEPPAPGE-PTPDGRL 774

RESULT 2
PCGN_RAT STANDARD: PRT: 1257 AA.
AC P53067;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurocan core protein precursor (245 kDa early postnatal core
glycoprotein) [Contains: 150 kDa adult core glycoprotein].
GN CSPG3 OR NCAN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OY NCBL_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=92406907; PubMed=1326557;
RA Rauch U., Karthikeyan L., Maurel P., Margolis R.U., Margolis R.K.;
RT "Cloning and primary structure of neurocan, a developmentally
regulated, aggregating chondroitin sulfate proteoglycan of brain.";
RL J. Biol. Chem. 267:19536-19547(1992).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=94230574; PubMed=7513709;
RA Friedlander D.R., Milev P., Karthikeyan L., Margolis R.K.,
Margolis R.U., Grumet M.;
RT "The neuronal chondroitin sulfate proteoglycan neurocan binds to the
neural cell adhesion molecules Ng-CAM/L1/NILE and N-CAM, and inhibits
neural cell adhesion and neurite outgrowth.";
RL J. Cell Biol. 125:669-680(1994).
CC -1- FUNCTION: MAY MODULATE NEURONAL ADHESION AND NEURITE GROWTH DURING
DEVELOPMENT BY BINDING TO NEURAL CELL ADHESION MOLECULES (NG-CAM
AND N-CAM). CHONDROITIN SULFATE PROTEOGLYCAN; BINDS TO HYALURONIC
ACID.
CC -1- TISSUE SPECIFICITY: EARLY POSTNATAL AND ADULT BRAIN; NOT EXPRESSED
IN KIDNEY, LUNG, LIVER AND MUSCLE.
CC -1- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO N-LINKED AND
O-LINKED OLIGOSACCHARIDES (BY SIMILARITY).
CC -1- PTM: TWO ISOFORMS WERE FOUND THAT PROBABLY ARISE BY PROTEOLYTIC
DEGRADATION. THE LARGE ISOFORM IS PREDOMINANT IN EARLY POSTNATAL

```

```

CC CC BRAIN, THE SMALL ISOFORM IS FOUND IN ADULT BRAIN.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 LINK DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M97161; AAC37679.1; -.
CC HSSP: P00740; IEDM.
CC Interpro: IPR000152; Asx_hydroxyl.
CC Interpro: IPR000561; EGF-like.
CC Interpro: IPR000742; EGF-1ike.
CC Interpro: IPR001881; EGF_Ca.
CC Interpro: IPR003599; Ig.
CC Interpro: IPR003006; Ig_MHC.
CC Interpro: IPR001304; Lectin_C.
CC Interpro: IPR000538; Link.
CC Interpro: IPR000436; Sushl_SCR_CCP.
CC Pfam: PF00008; EGF_2.
CC Pfam: PF00047; Ig_1.
CC Pfam: PF00059; Lectin_C; 1.
CC Pfam: PF00084; sushl_1.
CC Pfam: PF00193; Xlink_2.
CC ProDom: PD000918; Link_2.
CC SMART: SM00032; CCP_1.
CC SMART: SM00034; CLECT_1.
CC SMART: SM00179; EGF_CA_1.
CC SMART: SM00001; EGF_Like_1.
CC SMART: SM00409; IG_1.
CC SMART: SM00445; LINK_2.
CC PROSITE: PS00010; ASX_HYDROXYL_1.
CC PROSITE: PS00022; EGF_1; 3.
CC PROSITE: PS0186; EGF_2; 1.
CC PROSITE: PS0187; EGF_CA_1.
CC PROSITE: PS01241; LINK_2.
CC PROSITE: PS00615; C-TYPE-LECTIN_1; 1.
CC PROSITE: PS00411; C-TYPE-LECTIN_2; 1.
CC Glycoprotein: Hyaluronic acid; Proteoglycan; Immunoglobulin domain;
CC EGF-like domain; Calcium; Repeat; Lectin; Sushl; Signal.
CC SIGNAL 1 22
CC CHAIN 23 1257 NEUROCAN CORE PROTEIN.
CC DOMAIN 639 1257 150 KDA ADULT CORE GLYCOPROTEIN.
CC DOMAIN 51 146 IG-LIKE V-TYPE DOMAIN.
CC DOMAIN 176 253 LINK 1.
CC DOMAIN 274 355 LINK 2.
CC DOMAIN 949 985 EGF-LIKE 1.
CC DOMAIN 987 1023 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
CC DOMAIN 1025 1154 C-TYPE LECTIN.
CC DOMAIN 1155 1213 SUSHI.
CC DISULFID 38 139 BY SIMILARITY.
CC DISULFID 181 252 BY SIMILARITY.
CC DISULFID 205 226 BY SIMILARITY.
CC DISULFID 279 354 BY SIMILARITY.
CC DISULFID 303 324 BY SIMILARITY.
CC DISULFID 324 324 BY SIMILARITY.
CC DISULFID 953 964 BY SIMILARITY.
CC DISULFID 958 973 BY SIMILARITY.
CC DISULFID 975 984 BY SIMILARITY.
CC DISULFID 1029 1040 BY SIMILARITY.
CC DISULFID 1057 1149 BY SIMILARITY.
CC DISULFID 1125 1141 BY SIMILARITY.
CC DISULFID 1156 1199 BY SIMILARITY.
CC DISULFID 1185 1212 BY SIMILARITY.
CC CARBOHYD 121 121 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

FT CAROHD 737 737 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 944 944 O-LINKED (XYL. . .) (CHONDROITIN SULFATE).
FT CAROHD 967 967 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 1164 1164 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1257 AA; 135544 MW; 992B33DCFA19EEEB CRC64;

Query Match
Best Local Similarity 19.5%; Score 81.5; DB 1; Length 1257;
Matches 28; Conservative 10; Mismatches 25; Indels 17; Gaps 5;

QY 4 SLLPRPAVPPLRMQPG---PAHPVLSFLR-----PSMDLVSAFSLPLAPLS--PT 51
DB 610 SSIPEALSASVLSQASPGGSPDFIVAMLRAPKMLPLPHSTLVNVPSPILSPASP 669

QY 52 SVP-----ISPVSGRGPPDP 67
DB 670 SYPEQNAVPSVFG-AEDPE 688

RESULT 3
FXB2_MOUSE STANDARD: PRT: 428 AA.
AC 064733:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Forhead box protein B2 (Transcription factor FKX-4).
GN FOXB2 OR FKX4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97014266; PubMed=8861101;
RA Kaestner K.H., Schuetz G., Monaghan A.P.;
RT "Expression of the winged helix genes fkh-4 and fkh-5 defines domains
RL in the central nervous system.";
RN Mech. Dev. 55:221-230(1996).
RN [2]
RP SEQUENCE OF 4-114 FROM N.A.
RC STRAIN=129;
RX MEDLINE=93361500; PubMed=7689224;
RA Kaestner K.H., Lee K.H., Schloendorff J., Hiemisch H.;
RN Monaghan A.P., Schuetz G.;
RT "Six members of the mouse forhead gene family are developmentally
RL regulated.";
RN Proc. Natl. Acad. Sci. U.S.A. 90:7628-7631(1993).
RL [1]
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING EMBRYOGENESIS.
CC -1- SIMILARITY: CONTAINS 1 FORK-HEAD DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X92591; CA63335.1;
CC EMBL: X71942; CA50744.1;
CC HSSP: 063245; 2HPH.
CC TRANSPAC: T02442;
CC MGD: MGI:1347468; Foxb2.
CC InterPro: IPR001766; TF_Fork_head.
CC Pfam: PF00250; Fork_head; 1.
CC PRINTS: PR00053; FORKHEAD.
CC ProDom: PD000425; TF_Fork_head; 1.
CC SMART: SM00339; FH; 1.
CC PROSITE: PS00657; FORK_HEAD_1; 1.
CC PROSITE: PS00658; FORK_HEAD_2; 1.
CC PROSITE: PS50039; FORK_HEAD_3; 1.

```

```

KW DNA-binding; Nuclear protein; Transcription regulation.
FT DNA_BIND 12 103 FORK-HEAD.
FT DOMAIN 139 153 POLY-HIS.
FT DOMAIN 156 162 POLY-HIS.
FT DOMAIN 163 172 POLY-PRO.
FT DOMAIN 217 231 POLY-ALA.
FT DOMAIN 249 258 POLY-ALA.
FT DOMAIN 321 330 POLY-ALA.
FT DOMAIN 396 399 POLY-ALA.
SQ SEQUENCE 428 AA; 45170 MW; DB8ABEFD1E94AB10 CRC64;

Query Match
Best Local Similarity 19.1%; Score 80; DB 1; Length 428;
Matches 23; Conservative 10; Mismatches 25; Indels 8; Gaps 2;

QY 3 HSLLPRPAVPPLRMQPGPAHPVLSFLRPSMDLVSAFSLP-----LAPLSPTSPIS 56
DB 350 HSNQSEPAVPVPIK--PTPALPVTTLPPALSVPTASQQLPAPSTVCMAASPTAPILLE 407

QY 57 PVSGR 62
DB 408 PTAAGR 413

RESULT 4
RBP1_ARATH STANDARD: PRT: 1840 AA.
ID RBP1_ARATH
AC P18616; P31635; Q9S2S8;
DT 01-NOV-1990 (Rel. 16, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA-directed RNA polymerase II largest subunit (EC 2.7.7.6).
GN RPB205 OR RPII OR RPB1 OR AT4G35800 OR F4B14.70.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustoids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=91080867; PubMed=2259344;
RA Nawrath C., Scheil J., Koncz C.;
RN "Homologous domains of the largest subunit of eucaryotic RNA
RL polymerase II are conserved in plants.";
RN Mol. Gen. Genet. 223:65-75(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=91355869; PubMed=2103447;
RA Dietrich M.A., Prenger J.P., Guilfoyle T.J.;
RT "Analysis of the genes encoding the largest subunit of RNA polymerase
RL II in Arabidopsis and soybean.";
RN Plant Mol. Biol. 15:207-223(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083486; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Entlan K.-D., Terryn N.,
RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Schmidheini T.,
RA Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Gymonpretz B., Chuang Y.-J., Vandenbussche F.,
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoer E.,
RA Weltzinger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA Holzner E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA Woolman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA De Keyser A., Buyschaert C., Giesen J., Villarroel R., De Clercq R.,

```

RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
 RA Petter A., Rajandream M.A., Lyne M., Benes V., Reckmann S.,
 RA Botkova D., Bloeker H., Scharfe M., Grimm M., Lochner T.-H.,
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Aber S.,
 RA Gabel C., Fuchs M., Faltmann B., Granderath K., Dauner D., Herzl A.,
 RA Neumann S., Argilloy F., Vitale D., Liguori R., Piravandi E.,
 RA Massenet O., Ouilley F., Clabaud G., Muendlein A., Felber R.,
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Chedfor F., Cooke R., Berger C., Monfort A., Casarubia E.,
 RA Glibson T., Weber N., Vandenbol M., Barques N., Terol J., Torres A.,
 RA Perez-Perez A., Purrelle B., Bent E., Johnson S., Tacon D., Jesse C.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Biele C.,
 RA Frishman D., Haase D., Lemcke R., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Shee P., Cordes M., Abu-Threideh J.,
 RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Kramer J., Fulton J., Mardis E., Dante M., Pepin K., Hillier L.,
 RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., All J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
 RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 RA Granat S., Shohdy N., Hasegawa A., Hamed A., Lodhi M., Johnson A.,
 RA Chen E., Marra M., Martensen R., McCombie M.R.,
 RA "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 RT thaliana".
 RL Nature 402:769-777(1999).
 CC -I- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 SUBSTRATES.
 CC -I- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
 (RNA)(N).
 CC -I- SUBCELLULAR LOCATION: Nuclear.
 CC -I- PTM: THE TANDEN 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.
 CC THE PHOSPHORYLATION ACTIVATES POL2.
 CC -I- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES WERE
 FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
 PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
 III FOR 5S AND TRNA GENES.
 CC -I- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X52954; CA37130.1; -;
 DR EMBL: X52944; CA36735.1; -;
 DR EMBL: AL031986; CA21466.2; -;
 DR EMBL: AL161588; CAB81489.1; -;
 DR PIR: S12071; JDM01.
 DR PIR: S11960; JDM02.
 DR InterPro: IPR000684; RNA_POLIII_repeat.
 DR InterPro: IPR000722; RNA_POL_A.
 DR InterPro: IPR002879; RNA_POL_A2.
 DR Pfam: PF00623; RNA_POL_A; 1.
 DR Pfam: PF01854; RNA_POL_A2; 1.
 DR PROSITE: PS00115; RNA_POL_III_REPEAT; 23.
 KW Transcription; DNA-directed RNA polymerase; Transcription; Zinc; Repeat;
 KW DNA-binding; Nuclear protein; Phosphorylation; Zinc-finger.
 FT ZN_FING 66 79
 FT C2H2_TYPE (POTENTIAL).
 FT BY SIMILARITY.
 FT DNA_BIND 327 398
 FT 786 796
 FT DOMAIN 1330 1819
 FT 117 124
 FT CONFLICT 193 197
 FT IORRK -> NSKEE (IN REF. 2).
 FT CONFLICT 193 197

FT CONFLICT 298 298 A -> R (IN REF. 2).
 FT CONFLICT 303 303 E -> R (IN REF. 2).
 FT CONFLICT 401 417 KELVDYGRHPHPCKTGA -> VHLVFISFSET (IN REF.
 FT CONFLICT 428 428 1).
 FT CONFLICT 446 446 L -> S (IN REF. 1).
 FT CONFLICT 446 446 K -> RYLLSYSHSTHKKLFLEVMILMSQ (IN
 FT CONFLICT 739 739 REF. 2).
 FT CONFLICT 1062 1062 N -> D (IN REF. 1).
 FT CONFLICT 1062 1062 A -> R (IN REF. 1).
 FT CONFLICT 1089 1089 A -> P (IN REF. 2).
 FT CONFLICT 1720 1720 Y -> YSPISPSY (IN REF. 1).
 FT CONFLICT 1720 1720 Y -> YSPISPSY (IN REF. 1).
 SO SEQUENCE 1840 AA; 204688 MW; 8453621A0945C1B6 CRC64;
 Query Match 19.1%; Score 80; DB 1; Length 1840;
 Best Local Similarity 36.4%; Pred. No. 3.8;
 Matches 28; Conservative 3; Mismatches 36; Indels 10; Gaps 3;
 OY 7 PRPAAPVPLRMOP---GRAHPVLSFLRPSWDLVSAPYSLPLAPLSPTSPISPSVGR 62
 DB 1736 PSIAVSPSNARLSPASPSYSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSGA 1791
 OY 63 GPD--PDHVAVNSRY 77
 DB 1792 SPDYSPSAGYSPPLPGY 1808
 RESULT 5
 NT3_RAT
 ID NT3_RAT STANDARD; PRT; 2319 AA.
 AC Q9RI17;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Neurogenic locus notch homolog protein 3 precursor (Notch 3).
 GN NOTCH3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Haritunians T., Boulter J., Weimaster G., Schanen N.C.;
 RT "Rattus norvegicus mRNA for Notch 3."
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP FUNCTION.
 RX MEDLINE=21094508; PubMed=11182080;
 RA Tanigaki K., Nogaki F., Takahashi J., Tashiro K., Kurooka H.,
 RA Honjo T.;
 RT "Notch1 and Notch3 instructively restrict bFGF-responsive multipotent
 RT neural progenitor cells to an astroglial fate."
 RL Neuron 29:45-55(2001).
 RN [3]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=21331789; PubMed=11438922;
 RA Irvin D.K., Zurcher S.D., Nguyen T., Weinmaster G., Kornblum H.I.;
 RT "Expression patterns of Notch1, Notch2, and Notch3 suggest multiple
 RT functional roles for the Notch-DSL signaling system during brain
 RT development."
 RL J. Comp. Neurol. 436:167-181(2001).
 RL CL
 RL -I- FUNCTION: Functions as a receptor for membrane-bound ligands
 CC Jagged1, Jagged2 and Delta1 to regulate cell-fate determination.
 CC Upon ligand activation through the released notch intracellular
 CC domain (NICD) it forms a transcriptional activator complex with
 CC RBP-J kappa and activates genes of the enhancer of split locus.
 CC Affects the implementation of differentiation, proliferation and
 CC apoptotic programs (By similarity). Acts instructively to control
 CC the cell fate determination of CNS multipotent progenitor cells,
 CC resulting in astroglial induction and neuron/oligodendrocyte
 CC suppression.
 CC -I- SUBUNIT: Heterodimer of a C-terminal fragment N(7M) and a N-
 CC terminal fragment N(EC) which are probably linked by disulfide
 CC bonds (By similarity).
 CC


```
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Following
CC proteolytical processing NICD is translocated to the nucleus.
CC -1- TISSUE SPECIFICITY: Expressed in postnatal central nervous system
CC (CNS) germinal zones and, in early postnatal life, within
CC numerous cells throughout the CNS. It is more highly localized
CC to ventricular germinal zones.
CC -1- PTM: Synthesized in the endoplasmic reticulum as an inactive form
CC which is proteolytically cleaved by a furin-like convertase in the
CC trans-coli network before it reaches the plasma membrane to yield
CC an active, ligand-accessible form. Cleavage results in a C-
CC terminal fragment N(NTM) and a N-terminal fragment N(EC). Following
CC ligand binding, it is cleaved by TNF-alpha converting enzyme
CC (TACE) to yield a membrane-associated intermediate fragment called
CC notch extracellular truncation (NEXT). This fragment is then
CC cleaved by presenilin dependent gamma-secretase to release a
CC notch-derived peptide containing the intracellular domain (NICD)
CC from the membrane (By similarity).
CC -1- PTM: Phosphorylated (By similarity).
CC -1- SIMILARITY: BELONGS TO THE NOTCH FAMILY.
CC -1- SIMILARITY: CONTAINS 34 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF164486; AAD4653.2; -.
CC DR HSSP: P00740; 1EDM.
CC DR InterPro: IPR002110; ANK.
CC DR InterPro: IPR000152; Asx_hydroxyl.
CC DR InterPro: IPR000561; EGF-like.
CC DR InterPro: IPR000742; EGF-2.
CC DR InterPro: IPR001881; EGF-Ca.
CC DR InterPro: IPR001438; EGF-11.
CC DR InterPro: IPR002049; Lamlnln_EGF.
CC DR InterPro: IPR000800; Notch.
CC DR Pfam: PF00008; EGF_34.
CC DR Pfam: PF00023; ank; 6.
CC DR Pfam: PF00066; notch; 3.
CC DR PRINTS: PR00010; EGFBL00D.
CC DR PRINTS: PR00011; EGF_LAMININ.
CC DR PRINTS: PR01452; NOTCH.
CC DR SMART: SM00248; ANK; 5.
CC DR SMART: SM00179; EGF_CA; 20.
CC DR SMART: SM00001; EGF_Like; 12.
CC DR PROSITE: PS50088; ANK_REPEAT; 4.
CC DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
CC DR PROSITE: PS00010; ASX_HYDROXYL; 18.
CC DR PROSITE: PS00022; EGF-1; 33.
CC DR PROSITE: PS01186; EGF-2; 26.
CC DR PROSITE: PS01187; EGF_CA; 17.
CC DR Receptor: Transcription regulation: Activator; Differentiation;
CC Developmental protein; Repeat; ANK repeat; EGF-like domain;
CC Transmembrane; Glycoprotein; Signal; Phosphorylation.
CC KW SIGNAL.
CC FT CHAIN 1 40
CC FT CHAIN 41 2319
CC FT CHAIN 1631 2319
CC FT CHAIN 1664 2319
CC FT CHAIN 1646 1666
CC FT TRANSMEM 1667 2319
CC FT DOMAIN 41 79
CC FT DOMAIN 80 120
CC FT DOMAIN 121 158
CC FT DOMAIN 160 197
CC EGF-Like 4, CALCIUM-BINDING (POTENTIAL).
```

```
FT DOMAIN 199 236
FT DOMAIN 238 274
FT DOMAIN 276 314
FT DOMAIN 316 352
FT DOMAIN 353 391
FT DOMAIN 393 431
FT DOMAIN 433 469
FT DOMAIN 471 507
FT DOMAIN 509 545
FT DOMAIN 547 584
FT DOMAIN 584 620
FT DOMAIN 622 657
FT DOMAIN 659 695
FT DOMAIN 722 732
FT DOMAIN 736 772
FT DOMAIN 773 810
FT DOMAIN 812 849
FT DOMAIN 851 887
FT DOMAIN 889 924
FT DOMAIN 926 962
FT DOMAIN 964 1000
FT DOMAIN 1002 1036
FT DOMAIN 1038 1084
FT DOMAIN 1086 1122
FT DOMAIN 1124 1160
FT DOMAIN 1162 1205
FT DOMAIN 1207 1246
FT DOMAIN 1248 1289
FT DOMAIN 1291 1327
FT DOMAIN 1337 1375
FT DOMAIN 1384 1420
FT REPEAT 1425 1461
FT REPEAT 1461 1503
FT REPEAT 1469 1503
FT REPEAT 1840 1869
FT REPEAT 1873 1903
FT REPEAT 1907 1936
FT REPEAT 1940 1969
FT REPEAT 1973 1993
FT SITE 1573 1573
FT DISULFID 44 56
FT DISULFID 50 67
FT DISULFID 69 78
FT DISULFID 84 95
FT DISULFID 89 108
FT DISULFID 110 119
FT DISULFID 125 136
FT DISULFID 130 146
FT DISULFID 148 157
FT DISULFID 164 176
FT DISULFID 170 185
FT DISULFID 187 196
FT DISULFID 203 214
FT DISULFID 208 224
FT DISULFID 226 235
FT DISULFID 242 253
FT DISULFID 247 262
FT DISULFID 264 273
FT DISULFID 280 293
FT DISULFID 287 302
FT DISULFID 304 313
FT DISULFID 320 331
FT DISULFID 325 340
FT DISULFID 342 351
FT DISULFID 357 368
FT DISULFID 362 379
FT DISULFID 381 390
FT DISULFID 397 410
FT DISULFID 404 419
FT DISULFID 421 430
FT DISULFID 437 448
FT DISULFID 442 457
FT DISULFID 459 468
FT EGF-Like 5.
FT EGF-Like 6, CALCIUM-BINDING (POTENTIAL).
FT EGF-Like 7.
FT EGF-Like 8, CALCIUM-BINDING (POTENTIAL).
FT EGF-Like 9.
FT EGF-Like 10, CALCIUM-BINDING (POTENTIAL).
FT EGF-Like 11, CALCIUM-BINDING (POTENTIAL).
FT EGF-Like 12, CALCIUM-BINDING (POTENTIAL).
FT EGF-Like 13, CALCIUM-BINDING (POTENTIAL).
FT EGF-Like 14, CALCIUM-BINDING (POTENTIAL).
FT EGF-Like 15, CALCIUM-BINDING (POTENTIAL).
FT EGF-Like 16, CALCIUM-BINDING (POTENTIAL).
FT EGF-Like 17, CALCIUM-BINDING (POTENTIAL).
FT EGF-Like 18.
FT EGF-Like 19.
FT EGF-Like 20.
FT EGF-Like 21, CALCIUM-BINDING (POTENTIAL).
FT EGF-Like 22, CALCIUM-BINDING (POTENTIAL).
FT EGF-Like 23, CALCIUM-BINDING (POTENTIAL).
FT EGF-Like 24.
FT EGF-Like 25.
FT EGF-Like 26.
FT EGF-Like 27.
FT EGF-Like 28.
FT EGF-Like 29, CALCIUM-BINDING (POTENTIAL).
FT EGF-Like 30.
FT EGF-Like 31, CALCIUM-BINDING (POTENTIAL).
FT EGF-Like 32.
FT EGF-Like 33.
FT EGF-Like 34.
FT EGF-Like 35.
FT EGF-Like 36.
FT EGF-Like 37.
FT EGF-Like 38.
FT EGF-Like 39.
FT EGF-Like 40.
FT EGF-Like 41.
FT EGF-Like 42.
FT EGF-Like 43.
FT EGF-Like 44.
FT EGF-Like 45.
FT EGF-Like 46.
FT EGF-Like 47.
FT EGF-Like 48.
FT EGF-Like 49.
FT EGF-Like 50.
FT EGF-Like 51.
FT EGF-Like 52.
FT EGF-Like 53.
FT EGF-Like 54.
FT EGF-Like 55.
FT EGF-Like 56.
FT EGF-Like 57.
FT EGF-Like 58.
FT EGF-Like 59.
FT EGF-Like 60.
FT EGF-Like 61.
FT EGF-Like 62.
FT EGF-Like 63.
FT EGF-Like 64.
FT EGF-Like 65.
FT EGF-Like 66.
FT EGF-Like 67.
FT EGF-Like 68.
FT EGF-Like 69.
FT EGF-Like 70.
FT EGF-Like 71.
FT EGF-Like 72.
FT EGF-Like 73.
FT EGF-Like 74.
FT EGF-Like 75.
FT EGF-Like 76.
FT EGF-Like 77.
FT EGF-Like 78.
FT EGF-Like 79.
FT EGF-Like 80.
FT EGF-Like 81.
FT EGF-Like 82.
FT EGF-Like 83.
FT EGF-Like 84.
FT EGF-Like 85.
FT EGF-Like 86.
FT EGF-Like 87.
FT EGF-Like 88.
FT EGF-Like 89.
FT EGF-Like 90.
FT EGF-Like 91.
FT EGF-Like 92.
FT EGF-Like 93.
FT EGF-Like 94.
FT EGF-Like 95.
FT EGF-Like 96.
FT EGF-Like 97.
FT EGF-Like 98.
FT EGF-Like 99.
FT EGF-Like 100.
FT EGF-Like 101.
FT EGF-Like 102.
FT EGF-Like 103.
FT EGF-Like 104.
FT EGF-Like 105.
FT EGF-Like 106.
FT EGF-Like 107.
FT EGF-Like 108.
FT EGF-Like 109.
FT EGF-Like 110.
FT EGF-Like 111.
FT EGF-Like 112.
FT EGF-Like 113.
FT EGF-Like 114.
FT EGF-Like 115.
FT EGF-Like 116.
FT EGF-Like 117.
FT EGF-Like 118.
FT EGF-Like 119.
FT EGF-Like 120.
FT EGF-Like 121.
FT EGF-Like 122.
FT EGF-Like 123.
FT EGF-Like 124.
FT EGF-Like 125.
FT EGF-Like 126.
FT EGF-Like 127.
FT EGF-Like 128.
FT EGF-Like 129.
FT EGF-Like 130.
FT EGF-Like 131.
FT EGF-Like 132.
FT EGF-Like 133.
FT EGF-Like 134.
FT EGF-Like 135.
FT EGF-Like 136.
FT EGF-Like 137.
FT EGF-Like 138.
FT EGF-Like 139.
FT EGF-Like 140.
FT EGF-Like 141.
FT EGF-Like 142.
FT EGF-Like 143.
FT EGF-Like 144.
FT EGF-Like 145.
FT EGF-Like 146.
FT EGF-Like 147.
FT EGF-Like 148.
FT EGF-Like 149.
FT EGF-Like 150.
FT EGF-Like 151.
FT EGF-Like 152.
FT EGF-Like 153.
FT EGF-Like 154.
FT EGF-Like 155.
FT EGF-Like 156.
FT EGF-Like 157.
FT EGF-Like 158.
FT EGF-Like 159.
FT EGF-Like 160.
FT EGF-Like 161.
FT EGF-Like 162.
FT EGF-Like 163.
FT EGF-Like 164.
FT EGF-Like 165.
FT EGF-Like 166.
FT EGF-Like 167.
FT EGF-Like 168.
FT EGF-Like 169.
FT EGF-Like 170.
FT EGF-Like 171.
FT EGF-Like 172.
FT EGF-Like 173.
FT EGF-Like 174.
FT EGF-Like 175.
FT EGF-Like 176.
FT EGF-Like 177.
FT EGF-Like 178.
FT EGF-Like 179.
FT EGF-Like 180.
FT EGF-Like 181.
FT EGF-Like 182.
FT EGF-Like 183.
FT EGF-Like 184.
FT EGF-Like 185.
FT EGF-Like 186.
FT EGF-Like 187.
FT EGF-Like 188.
FT EGF-Like 189.
FT EGF-Like 190.
FT EGF-Like 191.
FT EGF-Like 192.
FT EGF-Like 193.
FT EGF-Like 194.
FT EGF-Like 195.
FT EGF-Like 196.
FT EGF-Like 197.
FT EGF-Like 198.
FT EGF-Like 199.
FT EGF-Like 200.
FT EGF-Like 201.
FT EGF-Like 202.
FT EGF-Like 203.
FT EGF-Like 204.
FT EGF-Like 205.
FT EGF-Like 206.
FT EGF-Like 207.
FT EGF-Like 208.
FT EGF-Like 209.
FT EGF-Like 210.
FT EGF-Like 211.
FT EGF-Like 212.
FT EGF-Like 213.
FT EGF-Like 214.
FT EGF-Like 215.
FT EGF-Like 216.
FT EGF-Like 217.
FT EGF-Like 218.
FT EGF-Like 219.
FT EGF-Like 220.
FT EGF-Like 221.
FT EGF-Like 222.
FT EGF-Like 223.
FT EGF-Like 224.
FT EGF-Like 225.
FT EGF-Like 226.
FT EGF-Like 227.
FT EGF-Like 228.
FT EGF-Like 229.
FT EGF-Like 230.
FT EGF-Like 231.
FT EGF-Like 232.
FT EGF-Like 233.
FT EGF-Like 234.
FT EGF-Like 235.
FT EGF-Like 236.
FT EGF-Like 237.
FT EGF-Like 238.
FT EGF-Like 239.
FT EGF-Like 240.
FT EGF-Like 241.
FT EGF-Like 242.
FT EGF-Like 243.
FT EGF-Like 244.
FT EGF-Like 245.
FT EGF-Like 246.
FT EGF-Like 247.
FT EGF-Like 248.
FT EGF-Like 249.
FT EGF-Like 250.
FT EGF-Like 251.
FT EGF-Like 252.
FT EGF-Like 253.
FT EGF-Like 254.
FT EGF-Like 255.
FT EGF-Like 256.
FT EGF-Like 257.
FT EGF-Like 258.
FT EGF-Like 259.
FT EGF-Like 260.
FT EGF-Like 261.
FT EGF-Like 262.
FT EGF-Like 263.
FT EGF-Like 264.
FT EGF-Like 265.
FT EGF-Like 266.
FT EGF-Like 267.
FT EGF-Like 268.
FT EGF-Like 269.
FT EGF-Like 270.
FT EGF-Like 271.
FT EGF-Like 272.
FT EGF-Like 273.
FT EGF-Like 274.
FT EGF-Like 275.
FT EGF-Like 276.
FT EGF-Like 277.
FT EGF-Like 278.
FT EGF-Like 279.
FT EGF-Like 280.
FT EGF-Like 281.
FT EGF-Like 282.
FT EGF-Like 283.
FT EGF-Like 284.
FT EGF-Like 285.
FT EGF-Like 286.
FT EGF-Like 287.
FT EGF-Like 288.
FT EGF-Like 289.
FT EGF-Like 290.
FT EGF-Like 291.
FT EGF-Like 292.
FT EGF-Like 293.
FT EGF-Like 294.
FT EGF-Like 295.
FT EGF-Like 296.
FT EGF-Like 297.
FT EGF-Like 298.
FT EGF-Like 299.
FT EGF-Like 300.
FT EGF-Like 301.
FT EGF-Like 302.
FT EGF-Like 303.
FT EGF-Like 304.
FT EGF-Like 305.
FT EGF-Like 306.
FT EGF-Like 307.
FT EGF-Like 308.
FT EGF-Like 309.
FT EGF-Like 310.
FT EGF-Like 311.
FT EGF-Like 312.
FT EGF-Like 313.
FT EGF-Like 314.
FT EGF-Like 315.
FT EGF-Like 316.
FT EGF-Like 317.
FT EGF-Like 318.
FT EGF-Like 319.
FT EGF-Like 320.
FT EGF-Like 321.
FT EGF-Like 322.
FT EGF-Like 323.
FT EGF-Like 324.
FT EGF-Like 325.
FT EGF-Like 326.
FT EGF-Like 327.
FT EGF-Like 328.
FT EGF-Like 329.
FT EGF-Like 330.
FT EGF-Like 331.
FT EGF-Like 332.
FT EGF-Like 333.
FT EGF-Like 334.
FT EGF-Like 335.
FT EGF-Like 336.
FT EGF-Like 337.
FT EGF-Like 338.
FT EGF-Like 339.
FT EGF-Like 340.
FT EGF-Like 341.
FT EGF-Like 342.
FT EGF-Like 343.
FT EGF-Like 344.
FT EGF-Like 345.
FT EGF-Like 346.
FT EGF-Like 347.
FT EGF-Like 348.
FT EGF-Like 349.
FT EGF-Like 350.
FT EGF-Like 351.
FT EGF-Like 352.
FT EGF-Like 353.
FT EGF-Like 354.
FT EGF-Like 355.
FT EGF-Like 356.
FT EGF-Like 357.
FT EGF-Like 358.
FT EGF-Like 359.
FT EGF-Like 360.
FT EGF-Like 361.
FT EGF-Like 362.
FT EGF-Like 363.
FT EGF-Like 364.
FT EGF-Like 365.
FT EGF-Like 366.
FT EGF-Like 367.
FT EGF-Like 368.
FT EGF-Like 369.
FT EGF-Like 370.
FT EGF-Like 371.
FT EGF-Like 372.
FT EGF-Like 373.
FT EGF-Like 374.
FT EGF-Like 375.
FT EGF-Like 376.
FT EGF-Like 377.
FT EGF-Like 378.
FT EGF-Like 379.
FT EGF-Like 380.
FT EGF-Like 381.
FT EGF-Like 382.
FT EGF-Like 383.
FT EGF-Like 384.
FT EGF-Like 385.
FT EGF-Like 386.
FT EGF-Like 387.
FT EGF-Like 388.
FT EGF-Like 389.
FT EGF-Like 390.
FT EGF-Like 391.
FT EGF-Like 392.
FT EGF-Like 393.
FT EGF-Like 394.
FT EGF-Like 395.
FT EGF-Like 396.
FT EGF-Like 397.
FT EGF-Like 398.
FT EGF-Like 399.
FT EGF-Like 400.
FT EGF-Like 401.
FT EGF-Like 402.
FT EGF-Like 403.
FT EGF-Like 404.
FT EGF-Like 405.
FT EGF-Like 406.
FT EGF-Like 407.
FT EGF-Like 408.
FT EGF-Like 409.
FT EGF-Like 410.
FT EGF-Like 411.
FT EGF-Like 412.
FT EGF-Like 413.
FT EGF-Like 414.
FT EGF-Like 415.
FT EGF-Like 416.
FT EGF-Like 417.
FT EGF-Like 418.
FT EGF-Like 419.
FT EGF-Like 420.
FT EGF-Like 421.
FT EGF-Like 422.
FT EGF-Like 423.
FT EGF-Like 424.
FT EGF-Like 425.
FT EGF-Like 426.
FT EGF-Like 427.
FT EGF-Like 428.
FT EGF-Like 429.
FT EGF-Like 430.
FT EGF-Like 431.
FT EGF-Like 432.
FT EGF-Like 433.
FT EGF-Like 434.
FT EGF-Like 435.
FT EGF-Like 436.
FT EGF-Like 437.
FT EGF-Like 438.
FT EGF-Like 439.
FT EGF-Like 440.
FT EGF-Like 441.
FT EGF-Like 442.
FT EGF-Like 443.
FT EGF-Like 444.
FT EGF-Like 445.
FT EGF-Like 446.
FT EGF-Like 447.
FT EGF-Like 448.
FT EGF-Like 449.
FT EGF-Like 450.
FT EGF-Like 451.
FT EGF-Like 452.
FT EGF-Like 453.
FT EGF-Like 454.
FT EGF-Like 455.
FT EGF-Like 456.
FT EGF-Like 457.
FT EGF-Like 458.
FT EGF-Like 459.
FT EGF-Like 460.
FT EGF-Like 461.
FT EGF-Like 462.
FT EGF-Like 463.
FT EGF-Like 464.
FT EGF-Like 465.
FT EGF-Like 466.
FT EGF-Like 467.
FT EGF-Like 468.
FT EGF-Like 469.
FT EGF-Like 470.
FT EGF-Like 471.
FT EGF-Like 472.
FT EGF-Like 473.
FT EGF-Like 474.
FT EGF-Like 475.
FT EGF-Like 476.
FT EGF-Like 477.
FT EGF-Like 478.
FT EGF-Like 479.
FT EGF-Like 480.
FT EGF-Like 481.
FT EGF-Like 482.
FT EGF-Like 483.
FT EGF-Like 484.
FT EGF-Like 485.
FT EGF-Like 486.
FT EGF-Like 487.
FT EGF-Like 488.
FT EGF-Like 489.
FT EGF-Like 490.
FT EGF-Like 491.
FT EGF-Like 492.
FT EGF-Like 493.
FT EGF-Like 494.
FT EGF-Like 495.
FT EGF-Like 496.
FT EGF-Like 497.
FT EGF-Like 498.
FT EGF-Like 499.
FT EGF-Like 500.
FT EGF-Like 501.
FT EGF-Like 502.
FT EGF-Like 503.
FT EGF-Like 504.
FT EGF-Like 505.
FT EGF-Like 506.
FT EGF-Like 507.
FT EGF-Like 508.
FT EGF-Like 509.
FT EGF-Like 510.
FT EGF-Like 511.
FT EGF-Like 512.
FT EGF-Like 513.
FT EGF-Like 514.
FT EGF-Like 515.
FT EGF-Like 516.
FT EGF-Like 517.
FT EGF-Like 518.
FT EGF-Like 519.
FT EGF-Like 520.
FT EGF-Like 521.
FT EGF-Like 522.
FT EGF-Like 523.
FT EGF-Like 524.
FT EGF-Like 525.
FT EGF-Like 526.
FT EGF-Like 527.
FT EGF-Like 528.
FT EGF-Like 529.
FT EGF-Like 530.
FT EGF-Like 531.
FT EGF-Like 532.
FT EGF-Like 533.
FT EGF-Like 534.
FT EGF-Like 535.
FT EGF-Like 536.
FT EGF-Like 537.
FT EGF-Like 538.
FT EGF-Like 539.
FT EGF-Like 540.
FT EGF-Like 541.
FT EGF-Like 542.
FT EGF-Like 543.
FT EGF-Like 544.
FT EGF-Like 545.
FT EGF-Like 546.
FT EGF-Like 547.
FT EGF-Like 548.
FT EGF-Like 549.
FT EGF-Like 550.
FT EGF-Like 551.
FT EGF-Like 552.
FT EGF-Like 553.
FT EGF-Like 554.
FT EGF-Like 555.
FT EGF-Like 556.
FT EGF-Like 557.
FT EGF-Like 558.
FT EGF-Like 559.
FT EGF-Like 560.
FT EGF-Like 561.
FT EGF-Like 562.
FT EGF-Like 563.
FT EGF-Like 564.
FT EGF-Like 565.
FT EGF-Like 566.
FT EGF-Like 567.
FT EGF-Like 568.
FT EGF-Like 569.
FT EGF-Like 570.
FT EGF-Like 571.
FT EGF-Like 572.
FT EGF-Like 573.
FT EGF-Like 574.
FT EGF-Like 575.
FT EGF-Like 576.
FT EGF-Like 577.
FT EGF-Like 578.
FT EGF-Like 579.
FT EGF-Like 580.
FT EGF-Like 581.
FT EGF-Like 582.
FT EGF-Like 583.
FT EGF-Like 584.
FT EGF-Like 585.
FT EGF-Like 586.
FT EGF-Like 587.
FT EGF-Like 588.
FT EGF-Like 589.
FT EGF-Like 590.
FT EGF-Like 591.
FT EGF-Like 592.
FT EGF-Like 593.
FT EGF-Like 594.
FT EGF-Like 595.
FT EGF-Like 596.
FT EGF-Like 597.
FT EGF-Like 598.
FT EGF-Like 599.
FT EGF-Like 600.
FT EGF-Like 601.
FT EGF-Like 602.
FT EGF-Like 603.
FT EGF-Like 604.
FT EGF-Like 605.
FT EGF-Like 606.
FT EGF-Like 607.
FT EGF-Like 608.
FT EGF-Like 609.
FT EGF-Like 610.
FT EGF-Like 611.
FT EGF-Like 612.
FT EGF-Like 613.
FT EGF-Like 614.
FT EGF-Like 615.
FT EGF-Like 616.
FT EGF-Like 617.
FT EGF-Like 618.
FT EGF-Like 619.
FT EGF-Like 620.
FT EGF-Like 621.
FT EGF-Like 622.
FT EGF-Like 623.
FT EGF-Like 624.
FT EGF-Like 625.
FT EGF-Like 626.
FT EGF-Like 627.
FT EGF-Like 628.
FT EGF-Like 629.
FT EGF-Like 630.
FT EGF-Like 631.
FT EGF-Like 632.
FT EGF-Like 633.
FT EGF-Like 634.
FT EGF-Like 635.
FT EGF-Like 636.
FT EGF-Like 637.
FT EGF-Like 638.
FT EGF-Like 639.
FT EGF-Like 640.
FT EGF-Like 641.
FT EGF-Like 642.
FT EGF-Like 643.
FT EGF-Like 644.
FT EGF-Like 645.
FT EGF-Like 646.
FT EGF-Like 647.
FT EGF-Like 648.
FT EGF-Like 649.
FT EGF-Like 650.
FT EGF-Like 651.
FT EGF-Like 652.
FT EGF-Like 653.
FT EGF-Like 654.
FT EGF-Like 655.
FT EGF-Like 656.
FT EGF-Like 657.
FT EGF-Like 658.
FT EGF-Like 659.
FT EGF-Like 660.
FT EGF-Like 661.
FT EGF-Like 662.
FT EGF-Like 663.
FT EGF-Like 664.
FT EGF-Like 665.
FT EGF-Like 666.
FT EGF-Like 667.
FT EGF-Like 668.
FT EGF-Like 669.
FT EGF-Like 670.
FT EGF-Like 671.
FT EGF-Like 672.
FT EGF-Like 673.
FT EGF-Like 674.
FT EGF-Like 675.
FT EGF-Like 676.
FT EGF-Like 677.
FT EGF-Like 678.
FT EGF-Like 679.
FT EGF-Like 680.
FT EGF-Like 681.
FT EGF-Like 682.
FT EGF-Like 683.
FT EGF-Like 684.
FT EGF-Like 685.
FT EGF-Like 686.
FT EGF-Like 687.
FT EGF-Like 688.
FT EGF-Like 689.
FT EGF-Like 690.
FT EGF-Like 691.
FT EGF-Like 692.
FT EGF-Like 693.
FT EGF-Like 694.
FT EGF-Like 695.
FT EGF-Like 696.
FT EGF-Like 697.
FT EGF-Like 698.
FT EGF-Like 699.
FT EGF-Like 700.
FT EGF-Like 701.
FT EGF-Like 702.
FT EGF-Like 703.
FT EGF-Like 704.
FT EGF-Like 705.
FT EGF-Like 706.
FT EGF-Like 707.
FT EGF-Like 708.
FT EGF-Like 709.
FT EGF-Like 710.
FT EGF-Like 711.
FT EGF-Like 712.
FT EGF-Like 713.
FT EGF-Like 714.
FT EGF-Like 715.
FT EGF-Like 716.
FT EGF-Like 717.
FT EGF-Like 718.
FT EGF-Like 719.
FT EGF-Like 720.
FT EGF-Like 721.
FT EGF-Like 722.
FT EGF-Like 723.
FT EGF-Like 724.
FT EGF-Like 725.
FT EGF-Like 726.
FT EGF-Like 727.
FT EGF-Like 728.
FT EGF-Like 729.
FT EGF-Like 730.
FT EGF-Like 731.
FT EGF-Like 732.
FT EGF-Like 733.
FT EGF-Like 734.
FT EGF-Like 735.
FT EGF-Like 736.
FT EGF-Like 737.
FT EGF-Like 738.
FT EGF-Like 739.
FT EGF-Like 740.
FT EGF-Like 741.
FT EGF-Like 742.
FT EGF-Like 743.
FT EGF-Like 744.
FT EGF-Like 745.
FT EGF-Like 746.
FT EGF-Like 747.
FT EGF-Like 748.
FT EGF-Like 749.
FT EGF-Like 750.
FT EGF-Like 751.
FT EGF-Like 752.
FT EGF-Like 753.
FT EGF-Like 754.
FT EGF-Like 755.
FT EGF-Like 756.
FT EGF-Like 757.
FT EGF-Like 758.
FT EGF-Like 759.
FT EGF-Like 760.
FT EGF-Like 761.
FT EGF-Like 762.
FT EGF-Like 763.
FT EGF-Like 764.
FT EGF-Like 765.
FT EGF-Like 766.
FT EGF-Like 767.
FT EGF-Like 768.
FT EGF-Like 769.
FT EGF-Like 770.
FT EGF-Like 771.
FT EGF-Like 772.
FT EGF-Like 773.
FT EGF-Like 774.
FT EGF-Like 775.
FT EGF-Like 776.
FT EGF-Like 777.
FT EGF-Like 778.
FT EGF-Like 779.
FT EGF-Like 780.
FT EGF-Like 781.
FT EGF-Like 782.
FT EGF-Like 783.
FT EGF-Like 784.
FT EGF-Like 785.
FT EGF-Like 786.
FT EGF-Like 787.
FT EGF-Like 788.
FT EGF-Like 789.
FT EGF-Like 790.
FT EGF-Like 791.
FT EGF-Like 792.
FT EGF-Like 793.
FT EGF-Like 794.
FT EGF-Like 795.
FT EGF-Like 796.
FT EGF-Like 797.
FT EGF-Like 798.
FT EGF-Like 799.
FT EGF-Like 800.
FT EGF-Like 801.
FT EGF-Like 802.
FT EGF-Like 803.
FT EGF-Like 804.
FT EGF-Like 805.
FT EGF-Like 806.
FT EGF-Like 807.
FT EGF-Like 808.
FT EGF-Like 809.
FT EGF-Like 810.
FT EGF-Like 811.
FT EGF-Like 812.
FT EGF-Like 813.
FT EGF-Like 814.
FT EGF-Like 815.
FT EGF-Like 816.
FT EGF-Like 817.
FT EGF-Like 818.
FT EGF-Like 819.
FT EGF-Like 820.
FT EGF-Like 821.
FT EGF-Like 822.
FT EGF-Like 823.
FT EGF-Like 824.
FT EGF-Like 825.
FT EGF-Like 826.
FT EGF-Like 827.
FT EGF-Like 828.
FT EGF-Like 829.
FT EGF-Like 830.
FT EGF-Like 831.
FT EGF-Like 832.
FT EGF-Like 833.
FT EGF-Like 834.
FT EGF-Like 835.
FT EGF-Like 836.
FT EGF-Like 837.
FT EGF-Like 838.
FT EGF-Like 839.
FT EGF-Like 840.
FT EGF-Like 841.
FT EGF-Like 842.
FT EGF-Like 843.
FT EGF-Like 844.
FT EGF-Like 845.
FT EGF-Like 846.
FT EGF-Like 847.
FT EGF-Like 848.
FT EGF-Like 849.
FT EGF-Like 850.
FT EGF-Like 851.
FT EGF-Like 852.
FT EGF-Like 853.
FT EGF-Like 854.
FT EGF-Like 855.
FT EGF-Like 856.
FT EGF-Like 857.
FT EGF-Like 858.
FT EGF-Like 859.
FT EGF-Like 860.
FT EGF-Like 861.
FT EGF-Like 862.
FT EGF-Like 863.
FT EGF-Like 864.
FT EGF-Like 865.
FT EGF-Like 866.
FT EGF-Like 867.
FT EGF-Like 868.
FT EGF-Like 869.
FT EGF-Like 870.
FT EGF-Like 871.
FT EGF-Like 872.
FT EGF-Like 873.
FT EGF-Like 874.
FT EGF-Like 875.
FT EGF-Like 876.
FT EGF-Like 877.
FT EGF-Like 878.
FT EGF-Like 879.
FT EGF-Like 880.
FT EGF-Like 881.
FT EGF-Like 882.
FT EGF-Like 883.
FT EGF-Like 884.
FT EGF-Like 885.
FT EGF-Like 886.
FT EGF-Like 887.
FT EGF-Like 888.
FT EGF-Like 889.
FT EGF-Like 890.
FT EGF-Like 891.
FT EGF-Like 892.
FT EGF-Like 893.
FT EGF-Like 894.
FT EGF-Like 895.
FT EGF-Like 896.
FT EGF-Like 897.
FT EGF-Like 898.
FT EGF-Like 899.
FT EGF-Like 900.
FT EGF-Like 901.
FT EGF-Like 902.
FT EGF-Like 903.
FT EGF-Like 904.
FT EGF-Like 905.
FT EGF-Like 906.
FT EGF-Like 907.
FT EGF-Like 908.
FT EGF-Like 909.
FT EGF-Like 910.
FT EGF-Like 911.
FT EGF-Like 912.
FT EGF-Like 
```

```

FT DISULFID 475 486 BY SIMILARITY.
FT DISULFID 480 495 BY SIMILARITY.
FT DISULFID 497 506 BY SIMILARITY.
FT DISULFID 513 524 BY SIMILARITY.
FT DISULFID 518 533 BY SIMILARITY.
FT DISULFID 535 544 BY SIMILARITY.
FT DISULFID 551 561 BY SIMILARITY.
FT DISULFID 556 570 BY SIMILARITY.
FT DISULFID 572 581 BY SIMILARITY.

Query Match 18.7%; Score 78; DB 1; Length 2319;
Best Local Similarity 39.3%; Pred. No. 7.5;
Matches 24; Conservative 2; Mismatches 23; Indels 12; Gaps 1;

QY 6 LPRPAVPVPLMQPGAPVLSFLRPSMDLVSAFYSFLAPLSPVSPISVSGRGP 65
DB 2163 LNPVAVPLDMARLPPAPPPSFL-----LPLAPGSQLNPATPVSPHERPP 2210

QY 66 P 66
DB 2211 P 2211

RESULT 6
YB8_YEAST STANDARD: PRT: 848 AA.
ID YB8_YEAST STANDARD: PRT: 848 AA.
AC P38266;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-DEC-1994 (Rel. 37, Last annotation update)
DE Hypothetical 92.8 kDa protein in PHO88-CMD1 intergenic region.
GN YBR108W OR YBR0901.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=9508357; PubMed=7900426;
RA Mannhaupt G., Stucka R., Ehmlé S., Vetter I., Feldmann H.;
RT Analysis of a 70 kb region on the right arm of yeast chromosome II."
RL Yeast 10:1363-1381(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X78993; CAA55611.1; -
CC EMBL: Z35977; CAH85063.1; -
CC PIR: S44688; S44688.
CC SGD: S0000312; YBR108W.
CC Hypothetical protein.
CC SEQUENCE 848 AA; 92762 MW; F33D371369FBAF97 CRC64;

Query Match 18.2%; Score 76; DB 1; Length 848;
Best Local Similarity 34.3%; Pred. No. 4.1;
Matches 23; Conservative 6; Mismatches 28; Indels 10; Gaps 3;

QY 13 PVLPMQPGAPV-----LSFLRPSMDLVSA---FYSFLAPLSPVSPISVSGR 62
DB 383 PVPVTRQPPQPPMOGNTIYPIPSLIDSTGSPHFVTPFPDPAAPKPKIDIPVDVSS 442

QY 63 GPDPDAH 69
DB 443 LPPPPH 449

RESULT 7

```

```

PER2_HUMAN STANDARD: PRT: 1255 AA.
ID PER2_HUMAN STANDARD: PRT: 1255 AA.
AC O15055;
DT 15-JUL-1999 (Rel. 38, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Period circadian protein 2.
GN PER2 OR KIAA0347.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97349984; PubMed=9205841;
RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro."
RL DNA Res. 4:141-150(1997).
RN [2]
RP REVISIONS TO C-TERMINUS.
RC TISSUE=Brain;
RA Nagase T., Ishikawa K.-I., Seki N., Nakajima D., Ohira M.,
RA Miyajima N., Kotani H., Nomura N., Ohara O.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP TISSUE EXPRESSION.
RX MEDLINE=98087121; PubMed=9427249;
RA Sheerman L.P., Zylka M.J., Weaver D.R., Kolakowski L.F. Jr.,
RA Reppert S.M.;
RT "Two period homologs: circadian expression and photic regulation in
RT the suprachiasmatic nuclei."
RL Neuron 19:1261-1269(1997).
CC -----
CC -1- FUNCTION: CIRCADIAN REGULATOR THAT MAY ACT AS A TRANSCRIPTION
CC FACTOR. BEHAVES AS A NEGATIVE ELEMENT IN CIRCADIAN TRANSCRIPTIONAL
CC LOOP. DOES NOT APPEAR TO BIND DNA, SUGGESTING INDIRECT
CC TRANSCRIPTIONAL INHIBITION. EXPRESSION OSCILLATES WITH A 24 HR
CC RHYTHM IN THE SUPRACHIASMATIC NUCLEUS (SCN) AND THE WHOLE EYES.
CC OSCILLATIONS ARE MAINTAINED UNDER CONSTANT DARKNESS AND ARE
CC RESPONSIVE TO CHANGES OF THE LIGHT/DARK CYCLES. THERE IS A 4 HOUR
CC TIME DELAY BETWEEN PER1 AND PER2 OSCILLATIONS. THE EXPRESSION
CC RHYTHMS APPEAR TO ORIGINATE FROM RETINA (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED. FOUND IN HEART, BRAIN,
CC PLACENTA, LUNG, LIVER, SKELETAL MUSCLE, KIDNEY AND PANCREAS.
CC -1- INDUCTION: BY LIGHT (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 2 PAS (PER-ARNT-SIM) DIMERIZATION DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB002345; BAA20804.2; ALT_INIT.
CC Genew: HGNC:8846; PER2.
CC MIM: 603426;
DR InterPro: IPR001610; PAC.
DR InterPro: IPR000014; PAS_domain.
DR SMART: SM00086; PAC; 1.
DR SMART: SM00091; PAS; 2.
DR PROSITE: PS00112; PAS; 1.
KW Transcription regulation; Nuclear protein; Repeat; Biological rhythms.
FT DOMAIN 109 146 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT DOMAIN 182 248 PAS 1.

```

```

FT DOMAIN 319 389 PAS.2.
FT DOMAIN 398 438 PAC.
FT DOMAIN 510 513 POLY-ARG.
FT DOMAIN 789 806 NUCLEAR LOCALIZATION SIGNAL.
FT DOMAIN 842 979 PRO-RICH.
SQ SEQUENCE 1255 AA: 136579 MW: 2AEF2C6BD4B6CB0 CRC64:

Query Match
Best Local Similarity 17.8%; Score 74.5; DB 1; Length 1255;
Matches 30; Conservative 10; Mismatches 26; Indels 51; Gaps 5;

OY 1 GTHSLPRF-----AAYVPLRMQ-----PGPAHVLSTLRPSMDLVSAFYSLPL 45
DB 863 GYVAAPPAHPHASFYVPAVDLQHFAVQPPFPAPLAPVAFMLPYSFGSTENLPQ 922
OY 46 A-PLSP-----TSVP-----ISPVSVGRGPP 66
DB 923 AEFPSOPFPSPHTLTSEMASASQEPFSPRTSIPROPACAPATRAVPPSAMGRASP 979

RESULT 8
YC18_HUMAN STANDARD; PRT; 864 AA.
AC 09ULK2:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein KIAA1218 (Fragment).
GN KIAA1218.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP TISSUE=Brain;
RC MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K.-I., Kikuno R., Hirotsawa M., Nomura N.,
RA Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 6:337-345(1999).
CC -!- SIMILARITY: BELONGS TO THE ATEXIN 7 FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB033044; BAA86532.1; -
KW Hypothetical protein.
FT NON_TER 1
FT DOMAIN 223 226 POLY-THR.
FT DOMAIN 552 664 POLY-SER.
FT DOMAIN 786 790 POLY-SER.
SQ SEQUENCE 864 AA: 92138 MW: DESC8E130E48DA23 CRC64:

Query Match
Best Local Similarity 17.6%; Score 73.5; DB 1; Length 864;
Matches 33; Conservative 4; Mismatches 21; Indels 35; Gaps 7;

OY 4 SLTPRPAA---VPVPLR-MQGPAPHVLSTLRPSMDLVSAFYSLPLASP-----T 51
DB 517 SPLSPAPAHITTPVPAVYLQ-----FSNP-----SAVY-LPSAPISSTLTSSYIMT 562
OY 52 SVPIPSVGVGRGPPDA-----HVAVNLS 75
DB 563 SAMLSNAAFVTSPPDSALMSHTTAPFPHVAATLS 595
```

```

RESULT 9
SM6B_RAT STANDARD; PRT; 887 AA.
AC 070141:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Semaphorin 6B precursor (Semaphorin Z) (Sema Z).
GN SEMA6B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mistat; TISSUE=Brain;
RC MEDLINE=98087397; PubMed=9427525;
RA Kikuchi K., Ishida H., Kimura T.;
RT "Molecular cloning of a novel member of semaphorin family genes,
RT semaphorin Z.";
RL Brain Res. Mol. Brain Res. 51:229-237(1997).
CC -!- FUNCTION: MAY PLAY A ROLE IN BOTH PERIPHERAL AND CENTRAL NERVOUS
CC SYSTEM DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DEVELOPMENTAL STAGE: DETECTED IN THE FIRST BRANCHIAL ARCH OF
CC EMBRYONIC DAY 11 (E11) EMBRYO, AND SUBSEQUENTLY IN THE MYOTOMES
CC AND THE DORSAL ROOT GANGLIA IN DEVELOPING SOMITES FROM E11.5
CC THROUGH E13.5, BUT NOT IN THE BRAIN. HOWEVER, AT E15, 18, 21 AND
CC P0, SEMA2 WAS HIGHLY EXPRESSED IN THE BRAIN.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: CONAINS 1 SEMA DOMAIN.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB000776; BAA25687.1; -
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR001627; Sema.
DR Pfam: PF01403; Sema; 1.
DR SMART: SM00423; PST; 1.
KW Signal, Transmembrane; Multigene family; Neurogenesis; Glycoprotein;
KW developmental protein.
FT SIGNAL 1 26
FT CHAIN 27 887 SEMAPHORIN 6B.
FT DOMAIN 27 605 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 606 626 POTENTIAL.
FT DOMAIN 627 887 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 239 549 SEMA.
FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 887 AA: 95752 MW: 09543F3F202CD301 CRC64:

Query Match
Best Local Similarity 17.6%; Score 73.5; DB 1; Length 887;
Matches 27; Conservative 6; Mismatches 31; Indels 11; Gaps 5;

OY 1 GTH-----SLTPPAVVPVPLRMQGPAPHVLSTLRP-SMD-----LVSAYSLPLASPT 51
DB 699 GPHDLSGLTPPEQTPLPQKRLP-TTPPHAHALPRAWDHSHALLSASASTSLILLAH 757
OY 52 SVPIPSVGVGRGPP 65
```

DB 758 RAPEQPVPTESGPE 772

RESULT 10

ZAP3_HUMAN STANDARD: PRT: 1822 AA.

ID ZAP3_HUMAN PRT: 1822 AA.

AC P49750; P49752; Q9P1V7; (Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Nuclear protein ZAP3 (ZAP113).

GN ZAP3.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RA [1]

RP SEQUENCE FROM N.A.

RA Rowen L., Madan A., Qin S., Abbasi N., Baradaran L., Birditt B., Bloom S., Dors M., Dickhoff R., Fleetwood P., Harrison G., James R., Kaur A., Madan A., Owen M.P., Ratcliffe A., Shaffer T., Hood L., "Sequencing of human chromosome 14q24.3 region."

RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 539-847 AND 1397-1822 FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=95319502; PubMed=7596406;

RA Sherrington R., Rogeev E.I., Liang Y., Rogeeva E.A., Levesque G., Ikeda M., Chi H., Lin C., Li G., Holman K., Tsuda T., Mar L., Focini J., F., Bruni A.C., Montesi M.P., Sorbi S., Ralnero I., Ralness L., Nee L., Chumakov I., Pollen D., Brookes A., Sainsau P., Polinsky R.J., Masco W., da Silva H.A.R., Haines J.L., Ralness J.M., St George-Hyslop P.H., Roses A.D., Fraser P.E., "Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease."

RT Nature 375:754-760(1995).

CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).

CC -1- CAUTION: Ref.2 sequence differs from that shown due to a frameshift in position 1661.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

CC EMBL: AC007956; AAF61275.1; -

DR EMBL: L40403; AAC42008.1; ALT_FRAME.

DR EMBL: L40400; AAC42006.1; -

KW Nuclear protein.

FT DOMAIN 15 205 PRO-RICH.

FT DOMAIN 382 430 GLN-RICH.

FT DOMAIN 807 1209 ARG-RICH.

FT DOMAIN 1488 1577 ARG-RICH.

FT CONFLICT 621 621 P -> S (IN REF. 2).

FT CONFLICT 1404 1404 T -> I (IN REF. 2).

FT CONFLICT 1821 1821 K -> E (IN REF. 2).

SO SEQUENCE 1822 AA; 204947 MW; 8E6CB83FE540C7D2 CRC64;

Query Match 17.6%; Score 73.5; DB 1; Length 1822;

Best Local Similarity 35.0%; Pred. No. 16;

Matches 21; Conservative 7; Mismatches 23; Indels 9; Gaps 2;

OY 7 PPAAPVPLRLQMPGA--HPVLSELRSDVLSAFYSLPLAPISPTSPVSVGRG 63

DB 471 PPAPLLPVSNGSAPPTTYHPPLQSGAPSEGVNSK-----APLSKSLAPYSSSSDQG 524

RESULT 11

NK4R_HUMAN

ID NK4R_HUMAN STANDARD: PRT: 440 AA.

AC P30098:

DT 01-APR-1993 (Rel. 25, Last sequence update)

DT 01-APR-1993 (Rel. 25, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Neuromedin K receptor (NK4R) (Neurokinin B receptor) (NK-4 receptor) (NK-4R) (K1R) (Neurokinin 4 receptor) (NK4).

GN TACR3L OR TAC3RL OR TCAR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RA [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Placenta;

RX MEDLINE=92237319; PubMed=1315051;

RA Xie G.-X., Miyajima A., Goldstein A., "Expression cloning of cDNA encoding a seven-helix receptor from human placenta with affinity for opioid ligands."

RL Proc. Natl. Acad. Sci. U.S.A. 89:4124-4128(1992).

RN [2]

RP CHARACTERIZATION.

RX MEDLINE=97103087; PubMed=8947459;

RA Donaldson L.F., Haskell C.A., Hanley M.R., "Functional characterization by heterologous expression of a novel cloned tachykinin peptide receptor."

RL Biochem. J. 320:1-5(1996).

CC -1- FUNCTION: THIS IS A RECEPTOR FOR THE TACHYKININ NEUROPEPTIDE NEUROMEDIN K (NEUROKININ B). THE RANK ORDER OF AFFINITY OF THIS RECEPTOR TO TACHYKININS IS: NEUROMEDIN K > SUBSTANCE P.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- TISSUE SPECIFICITY: HIGHEST LEVELS IN SKELETAL MUSCLE, LIVER, LUNG AND HEART. LOW LEVELS IN PANCREAS.

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. HIGHEST TO OTHER TACHYKININS RECEPTORS.

CC -1- CAUTION: Was originally (Ref.1) thought to be a kappa-type opioid receptor.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

CC EMBL: M84605; AAA36395.1; -

DR PIR: A44081; A44081.

DR HSSP: P02699; 1F88.

DR InterPro: IPR000276; GPCR_Rhodpsn.

DR Pfam: PF00001; 7tm.1; 1.

DR PRINTS: PR00237; GPCRHHODPSN.

DR PROSITE: PS00237; G-PROTEIN_RECPEP_FL_1; 1.

DR PROSITE: PS50262; G-PROTEIN_RECPEP_FL_2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein; Lipoprotein; Palmitate.

FT DOMAIN 1 59 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 60 82 1 (POTENTIAL).

FT DOMAIN 83 92 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 93 114 2 (POTENTIAL).

FT DOMAIN 115 134 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 135 156 3 (POTENTIAL).

FT DOMAIN 157 176 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 177 197 4 (POTENTIAL).

FT DOMAIN 198 220 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 221 245 5 (POTENTIAL).

FT DOMAIN 246 274 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 275 296 6 (POTENTIAL).

FT DOMAIN 297 309 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 310 334 7 (POTENTIAL).

FT DOMAIN 335 440 CYTOPLASMIC (POTENTIAL).

FT DISULFD 133 208 BY SIMILARITY.


```

CC -1- SIMILARITY: CONTAINS 1 BROMODOMAIN.
CC -1- SIMILARITY: CONTAINS 1 Z2-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: S66385; AAB28651.1; -
CC TRANSFAC: T01318; -
CC MGD: MGI:1098280; Crebbp.
CC InterPro: IPR001487; Bromodomain.
CC InterPro: IPR003101; KIX.
CC InterPro: IPR000197; TAZ_finger.
CC InterPro: IPR000433; ZnF_Z2.
CC Pfam: PF00439; Bromodomain; 1.
CC Pfam: PF00569; Z2; 1.
CC Pfam: PF02135; zf-TAZ; 2.
CC Pfam: PF02172; KIX; 1.
CC PRINTS: PRO0503; BROMODOMAIN.
CC SMART: SM00297; BROMO; 1.
CC SMART: SM00291; ZnF_Z2; 1.
CC PROSITE: PS00633; BROMODOMAIN_1; 1.
CC PROSITE: PS00014; BROMODOMAIN_2; 1.
CC PROSITE: PS01357; ZF_Z2_1; 1.
CC PROSITE: PS0135; ZF_Z2_2; 1.
CC Transcription regulation; Nuclear protein; Activator; Bromodomain;
KW Zinc-finger.
KW
FT DOMAIN 1104 1176 BROMODOMAIN.
FT ZN_FING 1702 1745 Z2-TYPE.
FT DOMAIN 1062 1065 POLY-GLU.
FT DOMAIN 1556 1563 POLY-GLU.
FT DOMAIN 1944 1949 POLY-PRO.
FT DOMAIN 1968 1971 POLY-GLN.
FT DOMAIN 2082 2086 POLY-GLN.
FT DOMAIN 2200 2216 POLY-GLN.
FT DOMAIN 2296 2299 POLY-GLN.
SO SEQUENCE 2441 AA; 265474 MW; 0ABBD28C3112F419 CRC64;

Query Match 17.1%; Score 71.5; DB 1; Length 2441;
Best Local Similarity 36.5%; Pred. No. 34;
Matches 23; Conservative 4; Mismatches 33; Indels 3; Gaps 2;

QY 4 SLPRPAAVVPLRMQGPAPHVLSFLRPSWDLVSATYSLPLAPLSTSPISPVSGRG 63
DB 843 SQLPCRPVNTOSPLHPPPPPASTAGM-PSLQHPATGMPQPAPAPQ-PSTPVSSGQT 899

QY 64 PDP 66
DB 900 PTP 902

RESULT 14
YD53_SYNY3 STANDARD; PRT; 291 AA.
AC P42350;
ID YD53_SYNY3
DB 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein slr1353.
DE SLR1353.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93222488; PubMed=8467083;
RA Malakhov M.P., Wada H., Los D.A., Sakamoto T., Murata N.;
RT "Structure of a cyanobacterial gene encoding the 50S ribosomal
protein L9.";
```

```

RL Plant Mol. Biol. 21:913-918(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hikosawa M., Sugita M., Saito S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RT DNA Res. 3:109-136(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D10716; BAA38818.1; -
CC EMBL: D90912; BAA18174.1; -
KW Hypothetical protein; Complete proteome.
SO SEQUENCE 291 AA; 31283 MW; 5AB7E3DD03C36390 CRC64;

Query Match 16.9%; Score 70.5; DB 1; Length 291;
Best Local Similarity 28.8%; Pred. No. 4.5;
Matches 23; Conservative 16; Mismatches 30; Indels 11; Gaps 5;

QY 4 SLPRPAAVVPLRMQGPAPHVLSFLRPSWDLVSATYSLP-LAPLSP-----SVP 54
DB 145 ALAPEPTLPAPISPPSP-DPVLSEPTPPPPAMVNSFNQPESSAPLDSELDQDFARP 203

QY 55 ISPVSYGRGD-PDAHVAVN 73
DB 204 ELPLAVEARKDSPEPDMAVS 223

RESULT 15
DPOL_HPBHE
ID DPOL_HPBHE STANDARD; PRT; 788 AA.
AC P13846;
DB 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE P protein (includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-
DE directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)).
OS Heron hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.
OX NCBI_TaxID=28300;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88333160; PubMed=3418788;
RA Sprengel R., Kallea E.F., Will H.;
RT "Isolation and characterization of a hepatitis B virus endemic in
RT herons.";
RL J. Virol. 62:3832-3839(1988).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
CC + (DNA)(N).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

DR EMBL; M22056; AAA45738.1; -.
DR PIR; A30082; JDLH.
DR InterPro: IPR001462; DNAPol_viral_C.
DR InterPro: IPR000201; DNAPol_viral_N.
DR InterPro: IPR000477; RVise.
DR Pfam; PF00078; rvc; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR ProDom: PD000814; DNAPol_viral_C; 1.
DR Transferase; RNA-directed DNA polymerase; DNA-directed DNA polymerase;
KW Hydrolase; Nuclease; Endonuclease; DNA replication; DNA-binding.
SQ SEQUENCE 788 AA; 90070 MW; FB44F38F75EADF44 CRC64;

Query Match 16.9%; Score 70.5; DB 1; Length 788;
Best Local Similarity 40.0%; Pred. No. 13;
Matches 18; Conservative 4; Mismatches 12; Indels 11; Gaps 1;

QY 31 RPSMDLVSAFYSLPLAPLSPTSVPIIS-----PVSVGRGP 64
DB 444 RISLDLSQAFYHPLPLAPASSSRLAVSDGKQVYFRKAPMGVGLSP 488

Search completed: March 4, 2003, 12:34:57
Job time : 4.85542 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:31:26 ; Search time 8.7249 Seconds
(without alignments)
1865.663 Million cell updates/sec

Title: US-09-234-208b-1

Perfect score: 418
Sequence: 1 GTHSLPRPAAVPVLRLMP.....VGRGPDPAHVAVLNLSRYEG 79

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	406	97.1	419	4 Q9UK79	Q9UK79 homo sapien
2	99.5	23.8	309	12 Q8U2B4	Q8U2B4 grapevine f
3	85	20.3	327	11 Q9UJK6	Q9UJK6 mus musculu
4	84.5	20.2	434	5 Q9G015	Q9G015 theileria p
5	84	20.1	995	11 Q35615	Q35615 mus musculu
6	82.5	19.7	316	10 Q8R2Y1	Q8R2Y1 oryza sativ
7	82	19.6	122	12 Q91115	Q91115 hepatitis e
8	82	19.6	122	12 Q8V730	Q8V730 swine hepat
9	81.5	19.5	816	11 Q70474	Q70474 rattus norv
10	81	19.4	200	10 Q82761	Q82761 arabidopsis
11	81	19.4	1006	10 Q9LMO1	Q9LMO1 arabidopsis
12	78.5	18.8	941	12 Q9IMY0	Q9IMY0 cercopithec
13	78	18.7	122	12 Q9YLR0	Q9YLR0 hepatitis e
14	78	18.7	763	2 Q9XDH2	Q9XDH2 mycobacteri
15	77.5	18.5	487	5 Q9V197	Q9V197 drosophila
16	77.5	18.5	1212	16 Q9L1C8	Q9L1C8 streptomyce

17	77	18.4	122	12 Q36612	Q36612 swine hepat
18	77	18.4	503	5 Q17585	Q17585 caenorhabd
19	77	18.4	690	10 Q94D41	Q94D41 oryza sativ
20	77	18.4	1044	4 Q8W4X9	Q8W4X9 homo sapien
21	76.5	18.3	1110	5 Q09493	Q09493 caenorhabd
22	76	18.2	309	3 Q9P6R1	Q9P6R1 schizosacch
23	76	18.2	356	10 Q9SAK1	Q9SAK1 arabidopsis
24	76	18.2	1044	4 Q9NXC6	Q9NXC6 homo sapien
25	75	17.9	189	12 Q87034	Q87034 soybean dha
26	75	17.9	189	12 Q91Q07	Q91Q07 soybean dha
27	75	17.9	189	12 Q08404	Q08404 soybean dha
28	75	17.9	295	16 Q9KXQ7	Q9KXQ7 streptomyce
29	75	17.9	813	10 Q94LD6	Q94LD6 oryza sativ
30	75	17.9	1081	4 Q9HBF3	Q9HBF3 oryza sativ
31	75	17.9	1992	3 Q9PFT1	Q9PFT1 neosporea
32	74.5	17.8	598	16 Q8YKN7	Q8YKN7 mycobacteri
33	74.5	17.8	894	10 Q9M6S0	Q9M6S0 arabidopsis
34	74.5	17.8	894	10 Q9SUS3	Q9SUS3 arabidopsis
35	74	17.7	384	11 Q9D789	Q9D789 mus musculu
36	73.5	17.6	574	4 Q9BNQ8	Q9BNQ8 homo sapien
37	73.5	17.6	2321	12 Q9DGT6	Q9DGT6 turkey herp
38	73	17.5	123	12 Q9W1K2	Q9W1K2 hepatitis e
39	73	17.5	639	4 Q9H9M1	Q9H9M1 homo sapien
40	73	17.5	1180	5 Q9VWM2	Q9VWM2 drosophila
41	73	17.5	1664	4 Q9BZES	Q9BZES homo sapien
42	73	17.5	3503	5 Q24292	Q24292 drosophila
43	72.5	17.3	229	10 Q94H11	Q94H11 oryza sativ
44	72.5	17.3	518	16 Q373B8	Q373B8 mycobacteri
45	72.5	17.3	715	2 Q9F9V7	Q9F9V7 mycobacteri

ALIGNMENTS

ID	Q9UK79	PRELIMINARY:	PRT:	419 AA.
AC	Q9UK79:			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Herstatin.			
GN	HER-2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-99415951; PubMed-10485918;			
RA	Doherty J.K., Bond C., Jaridim A., Adelman J.P., Clinton G.M.;			
RT	"The HER-2/neu receptor tyrosine kinase gene encodes a secreted			
RT	autoinhibitor.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 96:10869-10874(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Doherty J.K., Clinton G.M., Adelman J.P., Evans A.J., Henner W.D.;			
RL	Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AF177761; AAD56009.2; -			
DR	InterPro: IPR000494; EGFR_Like-domain.			
DR	InterPro: IPR002174; Furin-Like.			
DR	Pfam: PF00757; Furin-Like; 1.			
DR	Pfam: PF01030; Recep_L-domain; 1.			
DR	SMART; SM00261; FU; 1.			
SO	SEQUENCE 419 AA; 45472 MW; FECLBE347E2D030C CRC64;			
Query Match	97.1%; Score 406; DB 4; Length 419;			
Best local Similarity	97.5%; Pred. No. 2.2e-35;			
Matches	77; Conservative 1; Mismatches 1; Indels 0; Gaps 0;			
OY	1 GTHSLPRPAAVPVLRLMPGPAHVAFLPSWDLVSAFYSLPLAPLPTSVTPSPVSV 60			
DB	341 GTHSLPRPAAVPVLRLMPGPAHVAFLPSWDLVSAFYSLPLAPLPTSVTPSPVSV 400			

DR	MED: MG1:10954400; Zfpml.
DR	InterPro: IPR001005; MYB_DNA_binding.
DR	InterPro: IPR000822; ZnF_C2H2.
DR	Pfam: PF00096; Zf-C2H2_5
DR	PRINTS: PR00048; ZINCfinger.
DR	SMART: SM00355; ZnF_C2H2; 5.
DR	PROSITE: PS00037; MYB_1; UNKNOWN_1.
DR	PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_2.
DR	PROSITE: PS0157; ZINC_FINGER_C2H2_2.
KW	DNA-binding; Nuclear protein; Zinc-finger
SQ	SEQUENCE 995 AA; 105983 MW; 293255B28151ECB8 CRC64;
<hr/>	
Query Match	20.1%; Score 84; DB 11; Length 995;
Best Local Similarity	33.3%; Pred. NO. 1.2;
Matches 23; Conservative 10; Mismatches 28; Indels 8; Gaps 2;	
OY	7 PPAAPVPLMOCGAPAVLFLPPSWDLVSATSLPLALSPPTSPVSGRGP-- 64 : : Db 766 PPAPAGPAPVPVPSFTALPSSRP-----GSASAGPAPALSPSPVDGPIIDLKRRR 820
OY	65 -DPAHAVAV 72 : db 821 QSDPATPAL 829

	RESULT 6			
08RZY1	ID	PRELIMINARY;	PRT:	316 AA.
AC	08RZY1			
DT	01-JUN-2002 (TREMBLrel. 21, Created)			
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	P0034C09.31 protein.			
OS	Oryza sativa (japonica cultivar-group).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Eihartoideae; Oryzeae; Oryza.			
OX	NCBI_TaxID=39947;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV. NIPPONBARE;			
RA	Sasaki T., Matsumoto T., Yamamoto K.;			
RT	"Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC			
RT	clone:P0034C09."			
RL	Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.			
DR	EMBL; AP003450; BAB84618.1; -			
SQ	SEQUENCE 316 AA; 34223 MW; 686ABF29A419C259 CRC64;			
	Query Match 19.7%; Score 82.5; DB 10; Length 316;			
	Best Local Similarity 37.9%; Pred. No. 0.51;			
	Matches 2; Conservative 8; Mismatches 26; Indels 7; Gaps 3;			
QY	4 SLDPRAVPVPLMQGPAPHPVLFLRPSMDLSAFSLPLAPLSPVSPISVGSG 63 : : : : : : : : : : : : : : : : : Db 54 TLSSPPIVVVPM-LLPSPTRPVVFSMQPHRDLYPA-----LPESP-QVPGSSLSTLA 106			
QY	64 PDPDAH 69 . .			
Dd	107 PGSTRH 112			
	RESULT 7			
091115	ID	PRELIMINARY;	PRT:	122 AA.
AC	091115			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
DE	ORF3 protein.			
OS	Hepatitis E virus.			
OC	Viruses; ssRNA positive-strand viruses, no DNA stage;			
OC	Hepaditis E-like viruses.			

OX	NCBI_TaxID=12461;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=JRAL;
RA	Mishiro S.;
RL	Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=JRAL;
RX	MEDLINE=21396683; PubMed=11504536;
RA	Takahashi K., Iwata K., Watanabe N., Hatahara T., Baba K.,
RA	Mishiro S.;
RT	"Full-genome nucleotide sequence of a hepatitis E virus strain that
RT	may be indigenous to Japan."
RL	Virology 287:9-12(2001).
DR	EMBL: AP003430; BAB63940.1; -
DR	InterPro: IPR003384; HEV_ORF2.
DR	Pfam: PF02444; HEV_ORF2_1.
SQ	SEQUENCE 122 AA; 12313 MW; D6C56A8E54C587DD CRC64;

[illegible]

Db 112 V 112

RESULT 9

ID 070474 PRELIMINARY; PRT; 816 AA.

AC 070474;

DT 01-AUG-1998 (TReMBLrel. 07, Created)

DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)

DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

DE Neurocan (Fragment).

OC Rattus norvegicus (Rat).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE-DAWLEY; TISSUE=HIPPOCAMPUS;

RA Zachmann-Brand B., Schaller H.C.;

RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF060879; AAC15766.1; -

DR HSSP: P00740; 1EDM.

DR InterPro: IPR000152; Asx_hydroxyl.

DR InterPro: IPR000561; EGF-like.

DR InterPro: IPR000742; EGF_2.

DR InterPro: IPR001881; EGF_Ca.

DR InterPro: IPR001438; EGF_II.

DR InterPro: IPR001304; Lectin_C.

DR InterPro: IPR000538; Link.

DR Pfam: PF00008; EGF_2.

DR Pfam: PF00059; Lectin_C; 1.

DR Pfam: PF00193; Xlink; 1.

DR PRINTS: PRO0010; EGFBL00D.

DR PRINTS: PRO1265; LINKMODULE.

DR ProDom: PD000918; Link; 1.

DR SMART: SM00179; EGF_CA; 1.

DR SMART: SM00001; EGF_Like; 1.

DR SMART: SM00445; Link; 1.

DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.

DR PROSITE: PS50041; C_TYPE_LLECTIN_2; 1.

DR PROSITE: PS00022; EGF_1; UNKNOWN_3.

DR PROSITE: PS01186; EGF_2; 1.

DR PROSITE: PS01187; EGF_CA; 1.

DR PROSITE: PS01241; Link; 1.

KW Calcium-binding; EGF-like domain; Glycoprotein; Repeat.

FT NON_TER 1

FT NON_TER 816

SQ SEQUENCE 816 AA; 85578 MW; 957F5917AD10616E CRC64;

Query Match 19.5%; Score 81.5; DB 11; Length 816;

Best Local Similarity 35.0%; Pred. No. 1.7;

Matches 28; Conservative 10; Mismatches 25; Indels 17; Gaps 5;

QY 4 SLLEPPAAVPLRMQPG---PAHPVLSFLR-----PSMDLVSAFYSLPLAPLS--PT 51

Db 342 SSIPSESLASVLSQASPDGSDPFIVAMLRAPKRLWLPSTLTVPNVSPILPSAPLPS 401

QY 52 SVP-----ISPVSYGRGDPD 67

Db 402 SYPEEQAVRVSFG-AEDPE 420

RESULT 10

ID 082761 PRELIMINARY; PRT; 200 AA.

AC 082761;

DT 01-NOV-1998 (TReMBLrel. 08, Created)

DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)

DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

DE Putative phytoeyanin (Putative blue copper-binding protein).

GN ATG31050.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RX MEDLINE-20083487; PubMed=10617197;

RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D., Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V., Buell C.R., Ketchum K.A., Lee J.J., Renning C.M., Koo H., Moffat K.S., Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R., RACopenhaver G.P., Preuss D., Nierman W.C., White O., Eissen J.A., Salberg S.L., Fraser C.M., Venter J.C.;

RT *Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.*

RL Nature 402:761-768(1999).

RL [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RA Lin X.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RA Rounsley S.D., Kaul S., Lin X., Ketchum K.A., Crosby M.L., Brandon R.C., Sykes S.M., Mason T.M., Kerlavage A.R., Adams M.D., Somerville C.R., Venter J.C.;

RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RA Town C.D., Kaul S.;

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AC005311; AAC63847.1; -

DR EMBL: AC004669; AAM14981.1; -

DR HSSP: P00303; 2CBP.

DR InterPro: IPR003245; Plectanin_Like.

DR Pfam: PF02298; Cu_bind_Like; 1.

DR ProDom: PD003122; Plectanin_Like; 1.

SQ SEQUENCE 200 AA; 21475 MW; E669011C997E349C CRC64;

Query Match 19.4%; Score 81; DB 10; Length 200;

Best Local Similarity 34.3%; Pred. No. 0.45;

Matches 23; Conservative 5; Mismatches 23; Indels 16; Gaps 2;

QY 9 PAAVPLRMQPGPAHPVLSFLRPSMDLVSAFYSLPLAPLSPTSPVSYGRGDPDA 68

Db 132 PVAAPV-----PGVPRPSSFSPTS-----QSPLAESPVNHAPVQYQMGSPAP 175

QY 69 HVAVNLS 75

Db 176 HSAASNS 182

RESULT 11

ID 091M01 PRELIMINARY; PRT; 1006 AA.

AC 091M01;

DT 01-OCT-2000 (TReMBLrel. 15, Created)

DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)

DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

DE F7H2.17 protein.

GN F7H2.17.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RA Liu S.X., Sakano H., Yu G., Etgu P., Lee J., Lenz C., Pham P.,

QY	60	VGRGPDPDA	68
Db	69	----PAPPA	73

QY	52	SVPISPV	59
		:: :	
Db	181	VVPVAPVA	188

Tue Mar 4 13:30:20 2003

us-09-234-208b-1.rspt

Page 7

Search completed: March 4, 2003, 12:35:58
job time : 10.7249 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:32:36 ; Search time 4.28313 Seconds
(without alignments)
542.689 Million cell updates/sec

Title: US-09-234-208b-1

Perfect score: 418
Sequence: 1 GTHSLPPRAVPVPLRMQP.....VGRGPPDAHVAVNLSRYEG 79

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PTOUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/Backfilest.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	418	100.0	79	4	US-09-630-155-1
2	418	100.0	419	4	US-09-630-155-2
3	81.5	19.5	1257	1	US-08-340-428B-49
4	77	18.4	122	4	US-09-462-606-65
5	72	17.2	440	4	US-08-430-286A-9
6	71.5	17.1	604	2	US-08-468-576B-12
7	71.5	17.1	604	2	US-08-468-579B-12
8	71.5	17.1	604	3	US-08-468-577B-12
9	71.5	17.1	2441	1	US-08-194-468-2
10	71.5	17.1	2441	3	US-08-961-739-2
11	71.5	17.1	2441	4	US-09-514-247A-8
12	69.5	16.6	344	4	US-09-147-236-11
13	69.5	16.6	432	1	US-08-615-170-21
14	69.5	16.6	433	1	US-08-615-170-19
15	68.5	16.4	115	4	US-09-461-697-58
16	68.5	16.4	132	4	US-09-461-697-54
17	68.5	16.4	139	4	US-09-461-697-52
18	68.5	16.4	159	4	US-09-461-697-48
19	68	16.3	2972	4	US-09-579-181-2
20	68	16.3	3118	4	US-09-579-181-1
21	67.5	16.1	123	3	US-08-840-316-3
22	67.5	16.1	123	3	US-08-478-507-9
23	67.5	16.1	123	4	US-08-809-523-3
24	67.5	16.1	123	4	US-09-128-275A-9
25	67.5	16.1	123	4	US-08-471-971-3
26	67.5	16.1	123	4	US-09-553-427-9
27	67.5	16.1	123	4	US-09-462-606-13

28	67.5	16.1	123	4	US-09-462-606-59	Sequence 59, Appl
29	67.5	16.1	123	4	US-09-462-606-60	Sequence 60, Appl
30	67.5	16.1	123	4	US-09-462-606-62	Sequence 62, Appl
31	67.5	16.1	123	4	US-09-462-606-63	Sequence 63, Appl
32	67.5	16.1	123	4	US-09-462-606-64	Sequence 64, Appl
33	67.5	16.1	123	4	US-09-402-776-3	Sequence 3, Appl1
34	67.5	16.1	123	5	PCR-US93-08849A-3	Sequence 3, Appl1
35	67.5	16.1	123	5	PCR-US93-08849-3	Sequence 3, Appl1
36	67.5	16.1	124	1	US-08-240-049B-19	Sequence 19, Appl
37	67.5	16.1	124	4	US-08-542-634-21	Sequence 21, Appl
38	67.5	16.1	124	4	US-08-477-232-21	Sequence 21, Appl
39	67.5	16.1	124	5	PCR-US95-13703-21	Sequence 21, Appl
40	67.5	16.1	136	4	US-08-259-451-5	Sequence 5, Appl1
41	67.5	16.1	433	4	US-08-259-451-3	Sequence 3, Appl1
42	67	16.0	174	4	US-08-818-112-143	Sequence 143, App
43	67	16.0	174	4	US-08-818-111-138	Sequence 138, App
44	67	16.0	174	4	US-09-056-556-143	Sequence 143, App
45	67	16.0	174	4	US-09-072-596-138	Sequence 138, App

ALIGNMENTS

```
RESULT 1
US-09-630-155-1
: Sequence 1, Application US/09630155
: Patent No. 6A14130
:
: GENERAL INFORMATION:
: APPLICANT: Doherty, Joni Kristin and Gall M. Clinton
: TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: ADDRESS: DAVIS WRIGHT TREMAINE LLP
: STREET: 1501 Fourth Avenue, 2600 Century Square
: CITY: Seattle
: STATE: Washington
: COUNTRY: U.S.A.
: ZIP: 98101
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: PC compatible
: OPERATING SYSTEM: Windows95
: SOFTWARE: Word
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/630,155
: FILING DATE: 16-Jan-2001
: CLASSIFICATION: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: Davison, Barry L.
: REGISTRATION NUMBER: 47,309
: REFERENCE/DOCKET NUMBER: 49321-10
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 206 628-7621
: TELEFAX: 206 628-7699
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 79
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: unknown
: MOLECULE TYPE: HER-2 ECD antagonist
: SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-630-155-1

Query Match      100.0%; Score 418; DB 4; Length 79;
Best Local Similarity 100.0%; Pred. No. 5.2e-42;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GTHSLPPRAVPVPLRMQPGPAHVLSFLRPSWDVSAFTSLPLAPSPVSPVSV 60
Db 1 GTHSLPPRAVPVPLRMQPGPAHVLSFLRPSWDVSAFTSLPLAPSPVSPVSV 60
Oy 61 GRGPPDAHVAVNLSRYEG 79
```

Db 61 GRGPPDAHVAVNLSRYEG 79

RESULT 2

US-09-630-155-2
Sequence 2, Application US/09630155
Patent No. 6414130

GENERAL INFORMATION:

APPLICANT: Doherty, Joni Kristin and Gail M. Clinton
TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVIS WRIGHT TREMAINE LLP
STREET: 1501 Fourth Avenue, 2600 Century Square
City: Seattle
STATE: Washington
COUNTRY: U.S.A.
ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: PC compatible
OPERATING SYSTEM: Windows95
SOFTWARE: Word

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/630,155
FILING DATE: 16-Jan-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Davison, Barry L.
REGISTRATION NUMBER: 47,309
REFERENCE/DOCKET NUMBER: 49321-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206 628-7621
TELEFAX: 206 628-7693

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 419
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: polypeptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-630-155-2

Query Match 100.0%: Score 418; DB 4; Length 419;
Best Local Similarity 100.0%: Pred. NO. 4.2e-41;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTHSLPRPAVPVPLRMQPGPAHPLSLFRSPWDLVSFAFVSLPLAPLSPTSPVSPVSV 60

Db 341 GTHSLPRPAVPVPLRMQPGPAHPLSLFRSPWDLVSFAFVSLPLAPLSPTSPVSPVSV 400

OY 61 GRGPPDAHVAVNLSRYEG 79

Db 401 GRGPPDAHVAVNLSRYEG 419

RESULT 3

US-08-340-428B-49
Sequence 49, Application US/08340428B
Patent No. 5648465

GENERAL INFORMATION:

APPLICANT: MARCOLIS, Richard U.
APPLICANT: RAUCH, Uwe
APPLICANT: MARCOLIS, Renee K.
TITLE OF INVENTION: CLONING, EXPRESSION AND USES FOR A
TITLE OF INVENTION: NEUROCAN AS A CHONDROITIN SULFATE PROTEOGLYCAN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street, N.W.
CITY: Washington

STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,428B
FILING DATE: 14 No. 5648465member 1994
CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/922,911
FILING DATE: 03 August 1992
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.

REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: Margolis=1A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528

INFORMATION FOR SEQ ID NO: 49:

SEQUENCE CHARACTERISTICS:
LENGTH: 1257 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-340-428B-49

Query Match 19.5%: Score 81.5; DB 1; Length 1257;
Best Local Similarity 35.0%: Pred. No. 0.38;
Matches 28; Conservative 10; Mismatches 25; Indels 17; Gaps 5;

OY 4 SLPRPAVPVPLRMQPG---PAHPVLSFLR-----PSMDLYSARFSLPLAPLS--PT 51

Db 610 SSIPEALSAVSLQASPDGSPDFIVAMLRAPKMLLPHTLVPNNVPIPLSPASPLPS 669

OY 52 SVP---ISPVSGRGPPD 67

Db 670 SVPEQAVPVSVFG-AEDPE 688

RESULT 4

US-09-462-606-65
Sequence 65, Application US/09462606
Patent No. 6432408

GENERAL INFORMATION:

APPLICANT: MENG, XIANG-JIN
APPLICANT: Emerson, Suzanne U.
APPLICANT: Purcell, Robert H.
TITLE OF INVENTION: A SWINE HEPATITIS E VIRUS AND USES THEREOF
FILE REFERENCE: 20264267US1
CURRENT APPLICATION NUMBER: US/09/462,606
CURRENT FILING DATE: 2000-06-12
PRIOR APPLICATION NUMBER: US 60/053069
PRIOR FILING DATE: 1997-07-18
PRIOR APPLICATION NUMBER: PCT/US98/14665
PRIOR FILING DATE: 1998-07-17

NUMBER OF SEQ ID NOS: 65
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 65
LENGTH: 122
TYPE: PRT
ORGANISM: Hepatitis E virus
US-09-462-606-65

Query Match 18.4%: Score 77; DB 4; Length 122;
Best Local Similarity 36.1%: Pred. No. 0.067;
Matches 22; Conservative 10; Mismatches 23; Indels 6; Gaps 3;

COUNTRY: USA
ZIP: 10591-5144
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: System 7.5
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,579B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,276
FILING DATE: 05-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/872,646
FILING DATE: 08-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/715,181
FILING DATE: 14-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/441,703
FILING DATE: 04-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/312,543
FILING DATE: 17-FEB-1989
ATTORNEY/AGENT INFORMATION:
NAME: Kurt G. Briscoe
REGISTRATION NUMBER: 33,141
REFERENCE/DOCKET NUMBER: MDI 251.5-KGB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914) 332-1700
TELEFAX: (914) 332-1844
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 604 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-468-579B-12

Query Match 17.1%; Score 71.5; DB 2; Length 604;
Best Local Similarity 34.8%; Pred. No. 2.2;
Matches 23; Conservative 13; Mismatches 27; Indels 3; Gaps 2;

QY 10 AAVPVLRQPGAPHVLSFLRPSMDLVSAFVSLPLAPLSPVSPISPV--SVGRGPPD 67
DB 421 AAOPLNLSRPKTAEPVKPTSTONLFPASKTSPVNLPKKSIP-SPIGSLGRSSLD 479

QY 68 AHVAVN 73
DB 480 ILSSLN 485

RESULT 8
US-08-468-577B-12
Sequence 12, Application US/08468577B
Patent No. 6001804
GENERAL INFORMATION:
APPLICANT: Rabin, Daniel
TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGENS
TITLE OF INVENTION: OBTAINED BY MOLECULAR CLONING
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sprung Kramer Schaefer & Briscoe
STREET: 660 White Plains Road
CITY: Tarrytown
STATE: New York
COUNTRY: USA
ZIP: 10591-5144
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: System 7.5

SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,577B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,276
FILING DATE: 05-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/872,646
FILING DATE: 08-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/715,181
FILING DATE: 14-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/441,703
FILING DATE: 04-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/312,543
FILING DATE: 17-FEB-1989
ATTORNEY/AGENT INFORMATION:
NAME: Kurt G. Briscoe
REGISTRATION NUMBER: 33,141
REFERENCE/DOCKET NUMBER: MDI 251.8-KGB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914) 332-1700
TELEFAX: (914) 332-1844
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 604 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-468-577B-12

Query Match 17.1%; Score 71.5; DB 3; Length 604;
Best Local Similarity 34.8%; Pred. No. 2.2;
Matches 23; Conservative 13; Mismatches 27; Indels 3; Gaps 2;

QY 10 AAVPVLRQPGAPHVLSFLRPSMDLVSAFVSLPLAPLSPVSPISPV--SVGRGPPD 67
DB 421 AAOPLNLSRPKTAEPVKPTSTONLFPASKTSPVNLPKKSIP-SPIGSLGRSSLD 479

QY 68 AHVAVN 73
DB 480 ILSSLN 485

RESULT 9
US-08-194-468-2
Sequence 2, Application US/08194468
Patent No. 5750336
GENERAL INFORMATION:
APPLICANT: Montminy, Marc R.
TITLE OF INVENTION: ASSAYS FOR THE IDENTIFICATION OF
TITLE OF INVENTION: COMPOUNDS WHICH INHIBIT ACTIVATION OF CAMP AND MITOGEN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/194,468
FILING DATE: 10-FEB-1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9672
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)-546-4737
TELEFAX: (619)-546-9392
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2441 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-194-468-2

Query Match 17.1%; Score 71.5; DB 1; Length 2441;
Best Local Similarity 36.5%; Pred. No. 13;

Matches 23; Conservative 4; Mismatches 33; Indels 3; Gaps 2;

QY 4 SLPRPAVPVPLRMQPGFAHPVLSFLRPSWDLVSFAFYSILPLAPLSPTSVIPSVSGRG 63

DB 843 SQLPCPPVTQSLHPTPPPASTAAGM--PSLQHPTRAPGMTPPQPAAPTO-PSIPVSSGQT 899

QY 64 PDP 66

DB 900 PTP 902

RESULT 10

US-08-961-739-2
Sequence 2, Application US/08961739A
Patent No. 6063583

GENERAL INFORMATION:

APPLICANT: Montminy, Marc R.
TITLE OF INVENTION: Methods for Treating Diabetes Mellitus

FILE REFERENCE: SALK1650-1

CURRENT APPLICATION NUMBER: US/08/961,739A

CURRENT FILING DATE: 1997-10-31

EARLIER APPLICATION NUMBER: US 194,468

EARLIER FILING DATE: 1994-02-10

NUMBER OF SEQ ID NOS: 4

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2

LENGTH: 2441

TYPE: PRT

ORGANISM: Mus

FEATURE:

NAME/KEY: VARIANT

LOCATION: (1)...(2441)

OTHER INFORMATION: Xaa = Any Amino Acid

US-08-961-739-2

Query Match 17.1%; Score 71.5; DB 3; Length 2441;
Best Local Similarity 36.5%; Pred. No. 13;

Matches 23; Conservative 4; Mismatches 33; Indels 3; Gaps 2;

QY 4 SLPRPAVPVPLRMQPGFAHPVLSFLRPSWDLVSFAFYSILPLAPLSPTSVIPSVSGRG 63

DB 843 SQLPCPPVTQSLHPTPPPASTAAGM--PSLQHPTRAPGMTPPQPAAPTO-PSIPVSSGQT 899

QY 64 PDP 66

DB 900 PTP 902

RESULT 11
US-09-514-247A-8
Sequence 8, Application US/09514247A
Patent No. 6365361

GENERAL INFORMATION:

APPLICANT: TANABE SEIYAKU CO. LTD.

APPLICANT: TANIGUCHI, Tomoyasu

APPLICANT: MIZUKAMI, Junko

TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND ANTAGONIST TO
FILE REFERENCE: TANIGUCHI-6
CURRENT APPLICATION NUMBER: US/09/514,247A
CURRENT FILING DATE: 2000-02-28
PRIOR APPLICATION NUMBER: PCT/JP98/03734
PRIOR FILING DATE: 1998-08-24
PRIOR APPLICATION NUMBER: JP231084/1997
PRIOR FILING DATE: 1997-08-27
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 2441
TYPE: PRT
ORGANISM: mouse
US-09-514-247A-8

Query Match 17.1%; Score 71.5; DB 4; Length 2441;
Best Local Similarity 36.5%; Pred. No. 13;

Matches 23; Conservative 4; Mismatches 33; Indels 3; Gaps 2;

QY 4 SLPRPAVPVPLRMQPGFAHPVLSFLRPSWDLVSFAFYSILPLAPLSPTSVIPSVSGRG 63

DB 843 SQLPCPPVTQSLHPTPPPASTAAGM--PSLQHPTRAPGMTPPQPAAPTO-PSIPVSSGQT 899

QY 64 PDP 66

DB 900 PTP 902

RESULT 12

US-09-147-236-11
Sequence 11, Application US/09147236A
Patent No. 6316251

GENERAL INFORMATION:

APPLICANT: TONOUCHI, Naoto

APPLICANT: TSUCHIDA, Takayasu

APPLICANT: YOSHINAGA, Fumihito

APPLICANT: TAHARA, Naoki

APPLICANT: HAYASHI, Takahisa

TITLE OF INVENTION: NOVEL GENE, GROUP OF GENES, AND NOVEL BETA-GLUCOSIDASE

FILE REFERENCE: 6537-011-0PCT

CURRENT APPLICATION NUMBER: US/09/147,236A

CURRENT FILING DATE: 1999-04-08

EARLIER APPLICATION NUMBER: PCT/JP97/03633

EARLIER FILING DATE: 1997-10-09

NUMBER OF SEQ ID NOS: 12

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 11

LENGTH: 344

TYPE: PRT

ORGANISM: Acetobacter xylinum

FEATURE:

OTHER INFORMATION: n at positions 15741 and 15767 may be a, g, c, or

OTHER INFORMATION: t

US-09-147-236-11

Query Match 16.6%; Score 69.5; DB 4; Length 344;
Best Local Similarity 31.0%; Pred. No. 1.9;

Matches 22; Conservative 7; Mismatches 19; Indels 23; Gaps 3;

QY 3 HSLPRPAVPVPLRMQPGFAHPVLSFL-----RPSWDLVSFAFYSILPLAPLSPTSVIP 55

DB 158 YAAAPQVATPPV--FQPAFAVAVAQAQPVQKQEPS-----LSPTPK 201

QY 56 SPVSVGRGPP 66

DB 202 PAVSFMADRP 112

RESULT 13
US-08-615-170-21
Sequence 21, Application US/08615170
Patent No. 5776776

```

GENERAL INFORMATION:
APPLICANT: ORDAHL, Charles P.
APPLICANT: AZAKIE, Anthony
APPLICANT: MAR, Janet H.
APPLICANT: FARRANCE, Iain K.G.
APPLICANT: HALL, Deborah E.
APPLICANT: STEWART, Alexandre F.R.
APPLICANT: LARKIN, Sarah B.
TITLE OF INVENTION: DTEF-1 ISOFORMS AND USES THEREOF
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourile and Crew
STREET: Stewart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/615,170
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01526
FILING DATE: 06-FEB-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/191,493
FILING DATE: 04-FEB-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 23070-053120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-615-170-21

Query Match      16.6%; Score 69.5; DB 1; Length 432;
Best Local Similarity 35.0%; Pred. No. 2.5;
Matches 21; Conservative 11; Mismatches 17; Indels 11; Gaps 4;

OY      4 SLPRPAVVPRLKMPGPAHPVLSFLRPSWDIVS-----AFYSUPLAPLSP--TSVPI 55
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 155 SAAPRWMSGPIP--GPGPSQDIKPPAQPAVPYPIQPPMPPSIASYE-PLAPLPASAVPV 211

RESULT 14
US-08-615-170-19
Sequence 19, Application US/08615170
Patent No. 5776776
GENERAL INFORMATION:
APPLICANT: ORDAHL, Charles P.
APPLICANT: AZAKIE, Anthony
APPLICANT: MAR, Janet H.
APPLICANT: FARRANCE, Iain K.G.
APPLICANT: HALL, Deborah E.
APPLICANT: STEWART, Alexandre F.R.
APPLICANT: LARKIN, Sarah B.
TITLE OF INVENTION: DTEF-1 ISOFORMS AND USES THEREOF
NUMBER OF SEQUENCES: 32
```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourile and Crew
STREET: Stewart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/615,170
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01526
FILING DATE: 06-FEB-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/191,493
FILING DATE: 04-FEB-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 23070-053120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 433 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-615-170-19

Query Match      16.6%; Score 69.5; DB 1; Length 433;
Best Local Similarity 35.0%; Pred. No. 2.5;
Matches 21; Conservative 11; Mismatches 17; Indels 11; Gaps 4;

OY      4 SLPRPAVVPRLKMPGPAHPVLSFLRPSWDIVS-----AFYSUPLAPLSP--TSVPI 55
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 156 SAAPRWMSGPIP--GPGPSQDIKPPAQPAVPYPIQPPMPPSIASYE-PLAPLPASAVPV 212

RESULT 15
US-09-461-697-58
Sequence 58, Application US/09461697
Patent No. 6277974
GENERAL INFORMATION:
APPLICANT: COGENT NEUROSCIENCE, Inc.
APPLICANT: LO, Donald C.
APPLICANT: BARNEY, Shawn
APPLICANT: THOMAS, Mary Beth
APPLICANT: PORTBURY, Stuart D.
APPLICANT: PURANAM, Kasturi
APPLICANT: KATZ, Lawrence C.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
FILE REFERENCE: 10001-005-999
CURRENT APPLICATION NUMBER: US/09/461,697
NUMBER OF SEQ ID NOS: 466
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 58
LENGTH: 115
TYPE: PRT
ORGANISM: Homo sapiens
```

Tue Mar 4 13:30:18 2003

Query Match	16.48;	Score 68.5;	DB 4;	Length 115;
Best Local Similarity	27.8%;	Pred No 0.61;		

	27/08	160	101	9	29	27	4
Matches	Conservative	Mismatches	Indels	Gaps			

QY 1 GTHSLRPRAVPVPLRMQP-----GRAH--PVLSFLRPSWDL-----VSAFY 411

Db 11 GCHCMSSRRDLPPLPHSEPGVLDCLGPHLLPLLSPGSPCWLGHLHPSLHPSPASASH 70

QY 42 SLPLAPLSPTSVPISPVS-----VGRC 63

Search completed: March 4, 2003, 12:37:08
Job time : 6.28313 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:36:06 ; Search time 3.33133 Seconds
(without alignments)
1000.035 Million cell updates/sec

Title: US-09-234-208b-1

Perfect score: 418

Sequence: 1 GTHSLIPRAAVPVLRLMQP.....VGRGPDPAHVAVNLRYREG 79

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188354 seqs, 42170167 residues

Total number of hits satisfying chosen parameters: 188354

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/PCT05_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83.5	20.0	888	US-10-036-041-35	Sequence 35, Appl
2	83.5	20.0	888	US-10-028-072-544	Sequence 544, App
3	83.5	20.0	888	US-10-035-855-35	Sequence 35, Appl
4	83.5	20.0	888	US-10-121-049-544	Sequence 544, App
5	83.5	20.0	888	US-10-123-904-544	Sequence 544, App
6	83.5	20.0	888	US-10-140-470-544	Sequence 544, App
7	83.5	20.0	888	US-09-931-836-35	Sequence 35, Appl
8	83.5	20.0	888	US-10-175-746-544	Sequence 544, App
9	83.5	20.0	888	US-10-176-918-544	Sequence 544, App
10	83.5	20.0	888	US-10-176-921-544	Sequence 544, App
11	83.5	20.0	888	US-10-036-214-35	Sequence 35, Appl
12	83.5	20.0	888	US-10-137-865-544	Sequence 544, App
13	83.5	20.0	888	US-10-140-474-544	Sequence 544, App
14	83.5	20.0	888	US-10-035-719-35	Sequence 35, Appl
15	83.5	20.0	888	US-10-142-431-544	Sequence 544, App
16	83.5	20.0	888	US-10-143-114-544	Sequence 544, App
17	83.5	20.0	888	US-10-140-002-544	Sequence 544, App
18	83.5	20.0	888	US-10-036-342-35	Sequence 35, Appl
19	74	17.7	616	US-09-925-300-1519	Sequence 1519, Ap

20	73	17.5	3503	9	US-10-108-605-237	Sequence 237, App
21	72	17.2	440	9	US-09-966-782A-7	Sequence 7, Appl
22	71.5	17.1	2441	12	US-10-109-886-8	Sequence 8, Appl
23	71	17.0	463	9	US-10-029-180-80	Sequence 80, Appl
24	69.5	16.6	802	10	US-09-823-240-2	Sequence 2, Appl
25	69	16.5	265	10	US-09-819-254-2	Sequence 2, Appl
26	69	16.5	265	10	US-09-998-598-2592	Sequence 2592, Ap
27	68.5	16.4	115	10	US-09-922-261-58	Sequence 58, Appl
28	68.5	16.4	132	10	US-09-922-261-52	Sequence 52, Appl
29	68.5	16.4	139	10	US-09-922-261-54	Sequence 54, Appl
30	68.5	16.4	159	10	US-09-922-261-48	Sequence 48, Appl
31	68	16.3	171	10	US-09-925-297-659	Sequence 659, App
32	68	16.3	3298	9	US-10-160-758-16	Sequence 16, Appl
33	67.5	16.1	124	10	US-09-769-066-21	Sequence 21, Appl
34	67.5	16.1	279	9	US-10-001-857-159	Sequence 159, App
35	67.5	16.1	609	10	US-09-796-338A-5	Sequence 5, Appl
36	67.5	16.1	609	10	US-09-782-980-74	Sequence 74, Appl
37	66.5	15.9	105	9	US-09-764-868-1235	Sequence 1235, Ap
38	66.5	15.9	435	10	US-09-749-728B-29	Sequence 29, Appl
39	66	15.8	119	10	US-09-864-761-36209	Sequence 36209, A
40	66	15.8	549	9	US-09-989-442-100	Sequence 100, App
41	66	15.8	744	9	US-09-764-868-667	Sequence 667, App
42	65.5	15.7	218	9	US-10-001-873-35	Sequence 35, Appl
43	65.5	15.7	1114	10	US-09-740-066-10	Sequence 10, Appl
44	64.5	15.4	124	10	US-09-769-066-22	Sequence 22, Appl
45	64.5	15.4	405	10	US-09-803-126-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1

US-10-036-041-35

Sequence 35, Application US/10036041

Publication No. US20020192751A1

GENERAL INFORMATION:

APPLICANT: Desnoyers, Luc

APPLICANT: Baton, Dan L.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James

APPLICANT: Stewart, Timothy A.

APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William I.

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P3030R1C8

CURRENT FILING DATE: 2001-12-26

PRIOR APPLICATION NUMBER: US/10/036,041

PRIOR FILING DATE: 1998-05-15

PRIOR APPLICATION NUMBER: 60/112514

PRIOR FILING DATE: 1998-12-15

PRIOR APPLICATION NUMBER: 60/113300

PRIOR FILING DATE: 1998-12-22

PRIOR APPLICATION NUMBER: 60/113430

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: 60/113605

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: 60/113621

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: 60/114140

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: 60/115552

PRIOR FILING DATE: 1999-01-12

PRIOR APPLICATION NUMBER: 60/116843

PRIOR FILING DATE: 1999-01-22

PRIOR APPLICATION NUMBER: 60/125774

PRIOR FILING DATE: 1999-03-23

PRIOR APPLICATION NUMBER: 60/125778

PRIOR FILING DATE: 1999-03-23

PRIOR APPLICATION NUMBER: 60/125826
PRIOR FILING DATE: 1999-03-24
PRIOR APPLICATION NUMBER: 60/127035
PRIOR FILING DATE: 1999-03-31
PRIOR APPLICATION NUMBER: 60/127706
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/129122
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: 60/130359
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/131270
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131272
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131291
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/132371
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/132379
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/132383
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/135750
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/138166
PRIOR FILING DATE: 1999-06-08
PRIOR APPLICATION NUMBER: 60/144791
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/146970
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 09/311832
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/380142
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/644848
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 09/747259
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: 09/816744
PRIOR FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 09/854208
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/854280
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/874503
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: 09/869599
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 09/908, 827
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: PCT/US99/10733
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30720
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: PCT/US00/05601
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/23522
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: PCT/US00/23328
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: PCT/US00/34956

PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: PCT/US01/17800
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: PCT/US01/19692
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: PCT/US01/21066
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: PCT/US01/21735
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 80
SEQ ID NO 35
LENGTH: 888
TYPE: PRT
ORGANISM: Homo Sapien
US-10-036-041-35
Query Match 20.0%; Score 83.5; DA 9; Length 888;
Best Local Similarity 35.4%; Pred. No. 0.91;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;
QY 1 GTH---SLPPRAVPPLRMQPGAPHVLSLRP-SWD---LVSIFYSLPLAPLSFT 51
Db 698 GPHDLSGLPPEOTPLPQKRLPTP-HPHPALGPRAMDGHPLPLPASASSLLILAPA 756
QY 52 SVSPISVSGRGDPDDAHV 70
Db 757 RAPEQPPAGE-PTPDGRL 774
RESULT 2
US-10-028-072-544
Sequence 544, Application US/10028072
Publication No. US20030004311A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Tanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang
TITLE OF INVENTION:
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/028, 072
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18

[illegible]

PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090538
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07

Query Match 20.0%; Score 83.5; DB 9; Length 888;
Best Local Similarity 35.4%; Pred. No. 0.91;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;

Oy 1 GHH---SLPRPAVPPRLRMQGPAPHVLSFLRP-SMD---LVSIFYSLPLASPT 51
Db 698 GPHDLSGLPPPEOTPLPKRLPTP-HPHPALGPRAMDHGHPULPASASSLLLLAPA 756

Oy 52 SVTSPVSVGRCPDPDAHV 70
Db 757 RAPEQPPAPAGE-PTPDGRL 774

RESULT 3
US-10-035-855-35
Sequence 35, Application US/10035855
Publication No. US20030008348A1
GENERAL INFORMATION:
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Stewart, Timothy A.
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3030R1C4
CURRENT APPLICATION NUMBER: US/10/035,855
CURRENT FILING DATE: 2001-12-26
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/112514
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113300
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113430
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/114140
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/115552
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/116843
PRIOR FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 60/125774
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/125778
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/125826
PRIOR FILING DATE: 1999-03-24
PRIOR APPLICATION NUMBER: 60/127035

PRIOR FILING DATE: 1999-03-31
PRIOR APPLICATION NUMBER: 60/12706
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/129122
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: 60/130359
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/131270
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131272
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131291
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/132371
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/132379
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/132383
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/135750
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/138166
PRIOR FILING DATE: 1999-06-08
PRIOR APPLICATION NUMBER: 60/144791
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/146970
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 09/311832
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/380142
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/644848
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 09/747259
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: 09/816744
PRIOR FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 09/854208
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/854280
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/874503
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: 09/869599
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 09/908,827
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: PCT/US99/10733
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30720
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: PCT/US00/05601
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/23522
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: PCT/US00/23328
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: PCT/US00/34956
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: 2001-02-28

;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P330R1C160
;; CURRENT APPLICATION NUMBER: US/10/140,470
;; PRIOR FILING DATE: 2002-05-06
;; PRIOR APPLICATION removed - See Palm or File Wrapper
;; NUMBER OF SEQ ID NOS: 550
;; SEQ ID NO 544
;; LENGTH: 888
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-140-470-544

Query Match 20.0%; Score 83.5; DB 9; Length 888;
Best Local Similarity 35.4%; Pred. No. 0.91;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;

Oy 1 GTH---SLPRPAVVPVLRMQGPRAHVLSFLRP-SWD---LVSAFYSLAPLSPT 51
Db 698 GPHDSDGLPTEQTPLPQKRLPFR-HPHRALGPRAMDGHPLLPASASSLLLLAPA 756
Oy 52 SVPISPVSGRGPPDANV 70
Db 757 RAPEQPPAPGE-PTPDGRL 774

RESULT 7
US-09-931-836-35
;; Sequence 35, Application US/09931836
;; Publication No. US20030027249A1
;; GENERAL INFORMATION:
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Eaton, Dan L.
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Gutney, Austin L.
;; APPLICANT: Pan, James
;; APPLICANT: Stewart, Timothy A.
;; APPLICANT: Watanabe, Colin K.
;; APPLICANT: Wood, William I.
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P3030R1C1
;; CURRENT APPLICATION NUMBER: US/09/931,836
;; PRIOR FILING DATE: 2001-08-16
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/112514
;; PRIOR FILING DATE: 1998-12-15
;; PRIOR APPLICATION NUMBER: 60/113300
;; PRIOR FILING DATE: 1998-12-22
;; PRIOR APPLICATION NUMBER: 60/113430
;; PRIOR FILING DATE: 1998-12-23
;; PRIOR APPLICATION NUMBER: 60/113605
;; PRIOR FILING DATE: 1998-12-23
;; PRIOR APPLICATION NUMBER: 60/113621
;; PRIOR FILING DATE: 1998-12-23
;; PRIOR APPLICATION NUMBER: 60/114140
;; PRIOR FILING DATE: 1998-12-23
;; PRIOR APPLICATION NUMBER: 60/115552
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/116843
;; PRIOR FILING DATE: 1999-01-22
;; PRIOR APPLICATION NUMBER: 60/125774
;; PRIOR FILING DATE: 1999-03-23
;; PRIOR APPLICATION NUMBER: 60/125778
;; PRIOR FILING DATE: 1999-03-23
;; PRIOR APPLICATION NUMBER: 60/125826
;; PRIOR FILING DATE: 1999-03-24
;; PRIOR APPLICATION NUMBER: 60/127035
;; PRIOR FILING DATE: 1999-03-31
;; PRIOR APPLICATION NUMBER: 60/127706

;; PRIOR FILING DATE: 1999-04-05
;; PRIOR APPLICATION NUMBER: 60/129122
;; PRIOR FILING DATE: 1999-04-13
;; PRIOR APPLICATION NUMBER: 60/130359
;; PRIOR FILING DATE: 1999-04-21
;; PRIOR APPLICATION NUMBER: 60/131270
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131272
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131291
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/132371
;; PRIOR FILING DATE: 1999-05-04
;; PRIOR APPLICATION NUMBER: 60/132379
;; PRIOR FILING DATE: 1999-05-04
;; PRIOR APPLICATION NUMBER: 60/132383
;; PRIOR FILING DATE: 1999-05-04
;; PRIOR APPLICATION NUMBER: 60/135750
;; PRIOR FILING DATE: 1999-05-25
;; PRIOR APPLICATION NUMBER: 60/138166
;; PRIOR FILING DATE: 1999-06-08
;; PRIOR APPLICATION NUMBER: 60/144791
;; PRIOR FILING DATE: 1999-07-20
;; PRIOR APPLICATION NUMBER: 60/146970
;; PRIOR FILING DATE: 1999-08-03
;; PRIOR APPLICATION NUMBER: 60/162506
;; PRIOR FILING DATE: 1999-10-29
;; PRIOR APPLICATION NUMBER: 09/311832
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: 09/380142
;; PRIOR FILING DATE: 1999-08-25
;; PRIOR APPLICATION NUMBER: 09/644848
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 09/747259
;; PRIOR FILING DATE: 2000-12-20
;; PRIOR APPLICATION NUMBER: 09/816744
;; PRIOR FILING DATE: 2001-03-22
;; PRIOR APPLICATION NUMBER: 09/854208
;; PRIOR FILING DATE: 2001-05-10
;; PRIOR APPLICATION NUMBER: 09/854280
;; PRIOR FILING DATE: 2001-05-10
;; PRIOR APPLICATION NUMBER: 09/874503
;; PRIOR FILING DATE: 2001-06-05
;; PRIOR APPLICATION NUMBER: 09/869599
;; PRIOR FILING DATE: 2001-06-29
;; PRIOR APPLICATION NUMBER: 09/908,827
;; PRIOR FILING DATE: 2001-07-18
;; PRIOR APPLICATION NUMBER: PCT/US99/10733
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: PCT/US99/28551
;; PRIOR FILING DATE: 1999-12-02
;; PRIOR APPLICATION NUMBER: PCT/US99/30720
;; PRIOR FILING DATE: 1999-12-22
;; PRIOR APPLICATION NUMBER: PCT/US00/05601
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: PCT/US00/05841
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: PCT/US00/14042
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: PCT/US00/15264
;; PRIOR FILING DATE: 2000-06-02
;; PRIOR APPLICATION NUMBER: PCT/US00/23522
;; PRIOR FILING DATE: 2000-08-23
;; PRIOR APPLICATION NUMBER: PCT/US00/23328
;; PRIOR FILING DATE: 2000-08-24
;; PRIOR APPLICATION NUMBER: PCT/US00/32678
;; PRIOR FILING DATE: 2000-12-01
;; PRIOR APPLICATION NUMBER: PCT/US00/34956
;; PRIOR FILING DATE: 2000-12-20
;; PRIOR APPLICATION NUMBER: PCT/US01/06520
;; PRIOR FILING DATE: 2001-08-28
;; PRIOR APPLICATION NUMBER: PCT/US01/17800
;; PRIOR FILING DATE: 2001-06-01

FILE REFERENCE: P330R1C288
CURRENT APPLICATION NUMBER: US/10/176,921
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 544
LENGTH: 888
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-921-544

Query Match 20.0%; Score 83.5; DB 9; Length 888;
Best Local Similarity 35.4%; Pred. No. 0.91;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;

QY I GTH-----SLLRPAVPPVLMQGPAPHPVLSFLRP-SWD-----LYSAFYSLPLAPLSPT 51
Db 698 GPHDLSGLTPPEQTPYQKRLETPR-HPHPALGPRAMDHGHPLLPASASSLLLLAPA 756
QY 52 SVPISPVSGKQPPDAHV 70
Db 757 RAPEGPAPGE-PTPDGRL 774

RESULT 11
US-10-036-214-35
Sequence 35, Application US/10036214
Publication No. US20030032061A1
GENERAL INFORMATION:
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Stewart, Timothy A.
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACTS ENCODING THE SAME
FILE REFERENCE: P3030R1C11
CURRENT APPLICATION NUMBER: US/10/036,214
CURRENT FILING DATE: 2001-12-26
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/112514
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113300
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113430
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/114140
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/115552
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/116843
PRIOR FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 60/125774
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/125778
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/125826
PRIOR FILING DATE: 1999-03-24
PRIOR APPLICATION NUMBER: 60/127035
PRIOR FILING DATE: 1999-03-31
PRIOR APPLICATION NUMBER: 60/127706
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/129122

PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: 60/130359
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/131270
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131272
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131291
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/132371
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/132379
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/132383
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/135750
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/138166
PRIOR FILING DATE: 1999-06-08
PRIOR APPLICATION NUMBER: 60/144791
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/146970
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 09/311832
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/380142
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/644848
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 09/747259
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: 09/816744
PRIOR FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 09/854208
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/854280
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/874503
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: 09/869599
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 09/908,827
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: PCT/US99/10733
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30720
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: PCT/US00/05601
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/23522
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: PCT/US00/23328
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: PCT/US00/34956
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: PCT/US01/17800
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: PCT/US01/19692
PRIOR FILING DATE: 2001-06-20

PRIOR APPLICATION NUMBER: 60/113430	PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113605	PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621	PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/114140	PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/115552	PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/116843	PRIOR FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 60/125774	PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/125778	PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/125826	PRIOR FILING DATE: 1999-03-24
PRIOR APPLICATION NUMBER: 60/127035	PRIOR FILING DATE: 1999-03-31
PRIOR APPLICATION NUMBER: 60/127706	PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/129122	PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: 60/130359	PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/131270	PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131272	PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131291	PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/132371	PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/132379	PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/132383	PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/135750	PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/138166	PRIOR FILING DATE: 1999-06-08
PRIOR APPLICATION NUMBER: 60/144191	PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/146970	PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/162506	PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 09/311832	PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/380142	PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/644848	PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 09/747259	PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: 09/816744	PRIOR FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 09/854208	PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/854280	PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/874503	PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: 09/869599	PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 09/908, 827	PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: PCT/US99/107333	PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: PCT/US99/285513	PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/307200	

```

1 PRIOR FILING DATE: 1999-12-22
2 PRIOR APPLICATION NUMBER: PCT/US00/05601
3 PRIOR FILING DATE: 2000-03-01
4 PRIOR APPLICATION NUMBER: PCT/US00/05841
5 PRIOR FILING DATE: 2000-03-02
6 PRIOR APPLICATION NUMBER: PCT/US00/14042
7 PRIOR FILING DATE: 2000-05-22
8 PRIOR APPLICATION NUMBER: PCT/US00/15264
9 PRIOR FILING DATE: 2000-06-02
10 PRIOR APPLICATION NUMBER: PCT/US00/23522
11 PRIOR FILING DATE: 2000-08-23
12 PRIOR APPLICATION NUMBER: PCT/US00/23328
13 PRIOR FILING DATE: 2000-08-24
14 PRIOR APPLICATION NUMBER: PCT/US00/32678
15 PRIOR FILING DATE: 2000-12-01
16 PRIOR APPLICATION NUMBER: PCT/US00/34956
17 PRIOR FILING DATE: 2000-12-20
18 PRIOR APPLICATION NUMBER: PCT/US01/06520
19 PRIOR FILING DATE: 2001-02-28
20 PRIOR APPLICATION NUMBER: PCT/US01/17800
21 PRIOR FILING DATE: 2001-06-01
22 PRIOR APPLICATION NUMBER: PCT/US01/19692
23 PRIOR FILING DATE: 2001-06-20
24 PRIOR APPLICATION NUMBER: PCT/US01/21066
25 PRIOR FILING DATE: 2001-06-29
26 PRIOR APPLICATION NUMBER: PCT/US01/21735
27 PRIOR FILING DATE: 2001-07-09
28 NUMBER OF SEQ ID NOS: 80
29 SEQ ID NO 35
30 LENGTH: 888
31 TYPE: PRT
32 ORGANISM: Homo Sapien
33 US-10-035-719-35

```

Query Match	20.0%;	Score 83.5;	DB 9;	Length 888;
Best Local Similarity	35.4%;	Pred. No. 0.91;		
Matches	28;	Conservative	7;	Mismatches 33;
				Indels 11;
				Gaps 5

```

QY      1  GTI----SLPRRAAVVRLKRGGRRAHVLSFLRP-SWD----LVSAFYSLPLARLSPT 51
          | | | | | : | : | | | | | | | | | | | | | | | | | | | |
DB     698  GRHDLDSGLRPTREQGTRLRQKYLPT-RPNRAHGRRAWDHGRRLLRASASSLLLLARA 756

```

```
QY. 52 SVPISPVSVGRGPPDAHV 70
      | | : | | | :
Db 757 RAPEQPPAPGE-PTPDGRL 774
```

```

RESULT 15
US-10-142-431-544
; Sequence 544, Application US/10142431
; Publication No. US20030036179A1
; GENERAL INFORMATION:

```

```

1  APPLICANT: Baker, Kevin P.
2  APPLICANT: Beresini, Maureen
3  APPLICANT: DeForge, Laura
4  APPLICANT: Desnoyers, Luc
5  APPLICANT: Filvaroff, Ellen
6  APPLICANT: Gao, Wei-Qiang
7  APPLICANT: Gerritsen, Mary E.
8  APPLICANT: Goddard, Audrey
9  APPLICANT: Godowski, Paul J.
10 APPLICANT: Gurney, Austin L.
11 APPLICANT: Sherwood, Steven
12 APPLICANT: Smith, Victoria
13 APPLICANT: Stewart, Timothy A.
14 APPLICANT: Tumas, Daniel
15 APPLICANT: Watanabe, Colin K
16 APPLICANT: Wood, William
17 APPLICANT: Zhang, Zemin
18 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
19 TITLE OF INVENTION: ACIDS ENCODING THE SAME
20 FILE REFERENCE: P3330RIC251
21 CURRENT APPLICATION NUMBER: US/10/142,431

```

```
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 544
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-142-431-544

Query Match      20.0%; Score 83.5; DB 9; Length 888;
Best Local Similarity 35.4%; Pred. No. 0.91;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;

QY      1  GTH-----SLPRPAAYVPLRMQGPAPVLSFLRP-SWD---LVSATFSLPLAPLSPT 51
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      698  GPHDLDSGLPLPDEQTPLPQKRLPTP-HPHPALGPPAWDHGHPLLPASASSSLILLAPA 756
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      52  SVPISPVSVGRGPPDPDAHV 70
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      757  RAPEQPPAPGE-PTPDGRL 774
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: March 4, 2003, 12:43:43
Job time : 4.33133 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:25:00 ; Search time 95.9157 Seconds
(without alignments)
582.095 Million cell updates/sec

Title: US-09-234-208b-2

Perfect score: 2287
Sequence: 1 MELALCRWGLLALLPPGA.....YGRGPDPAHVAVNSRYEG 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq.101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2181	95.4	419	22	AAE09213 Human p68HER-2 gen
2	2176	95.1	419	22	AAE09212 Human p68HER-2 gen
3	2174	95.1	419	22	AAE09205 Human p68HER-2 gen
4	2172	95.0	419	22	AAE09207 Human p68HER-2 gen
5	2172	95.0	419	22	AAE09209 Human p68HER-2 gen
6	2171	94.9	419	23	AAE20348 Human truncated HE
7	2170	94.9	419	22	AAE09203 Human p68HER-2 gen
8	2169	94.8	419	22	AAE09181 Human p68HER-2 gen
9	2169	94.8	419	22	AAE09208 Human p68HER-2 gen
10	2169	94.8	419	22	AAE09210 Human p68HER-2 gen

11	2168	94.8	419	22	AAE09206 Human p68HER-2 gen
12	2168	94.8	419	22	AAE09211 Human p68HER-2 gen
13	2167	94.8	419	22	AAE09204 Human p68HER-2 gen
14	2164	94.6	419	22	AAE09216 Human p68HER-2 gen
15	2161	94.5	419	22	AAE09200 Human p68HER-2 gen
16	2161	94.5	419	22	AAE09202 Human p68HER-2 gen
17	2158.5	94.4	420	21	AAV97240 Truncated HER-2, p
18	2157	94.3	419	22	AAE09183 Human p68HER-2 gen
19	2155	94.2	419	22	AAE09214 Human p68HER-2 gen
20	2155	94.2	419	22	AAE09215 Human p68HER-2 gen
21	1878	82.1	645	22	AAE09208 Human ErbB2 oncopr
22	1878	82.1	645	22	AAE09209 Human ErbB2 extrac
23	1878	82.1	653	21	AAE21203 Extracellular HER-
24	1878	82.1	653	23	AAE21145 Human HER-2/neu on
25	1878	82.1	712	23	AAE21204 Human HER-2/neu fu
26	1878	82.1	712	23	AAE21149 Her-2/neu extracel
27	1878	82.1	782	18	AAE19764 Her-2-GM-CSF immuno
28	1878	82.1	919	21	AAE21203 Human HER-2/neu fu
29	1878	82.1	919	23	AAE21148 Human HER-2/neu pr
30	1878	82.1	1200	21	AAE21208 Human HER-2/neu pr
31	1878	82.1	1255	17	AAE01111 Human HER-2/neu pr
32	1878	82.1	1255	20	AAE2406 Human HER-2/neu on
33	1878	82.1	1255	21	AAE21198 Human HER-2/neu pr
34	1878	82.1	1255	21	AAE2406 Human HER-2/neu pr
35	1878	82.1	1255	21	AAE2406 Human HER-2/neu pr
36	1878	82.1	1255	22	AAE21203 Human HER-2/neu pr
37	1878	82.1	1255	22	AAE21203 Human HER-2/neu pr
38	1878	82.1	1255	22	AAE21203 Human HER-2/neu pr
39	1878	82.1	1255	22	AAE21203 Human HER-2/neu pr
40	1878	82.1	1255	22	AAE21203 Human HER-2/neu pr
41	1878	82.1	1255	23	AAE20479 Human HER-2/neu pr
42	1878	82.1	1255	23	AAE20479 Human HER-2/neu pr
43	1878	82.1	1255	23	AAE20479 Human HER-2/neu pr
44	1878	82.1	1255	23	AAE20479 Human HER-2/neu pr
45	1860	81.3	1433	14	AAE39568 Sequence of c-erbB

ALIGNMENTS

RESULT 1	
AAE09213	
ID	AAE09213 standard; Protein; 419 AA.
XX	
AC	AAE09213;
DT	15-NOV-2001 (first entry)
XX	
DE	Human p68HER-2 generic protein variant 11.
XX	
KW	HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW	Solid tumour; cancer; polymorphism; cytostatic; gene therapy;
KW	p68HER-2; ECDIIIA; variant.
XX	
OS	Homo sapiens.
XX	
Key	Location/Qualifiers
FT	1..340 /note= "Identical to N-terminal region of p185HER-2"
FT	341..419 /label= "ECDIIIA variant"
FT	124 /note= "Extracellular domain IIIA variant"
FT	125 /note= "Represented as Agn in the parent sequence shown in the specification"
FT	342 /note= "Represented as Agn in the parent sequence shown in the specification"
FT	345 /label= Unknown
FT	346 /label= Unknown
FT	Misc-difference 346

```

FT      /note="p68HER-2 generic sequence (AAE09181) Xaa
FT      substituted with Leu"
FT      Misc-difference 356
FT      /Label= Unknown
FT      Misc-difference 358
FT      /Label= Unknown
FT      Misc-difference 361
FT      /Label= Unknown
FT      Misc-difference 376
FT      /Label= Unknown
FT      Misc-difference 394
FT      /Label= Unknown
FT      Misc-difference 404
FT      /Label= Unknown
FT      Misc-difference 413
FT      /note="p68HER-2 generic sequence (AAE09181) Xaa
FT      substituted with Asn"
FT      WO200161356-A1.
FT      23-AUG-2001.
FT      16-FEB-2001; 2001WO-US05327.
FT      16-FEB-2000; 2000US-0506079.
FT      (UYOR-) UNIV OREGON HEALTH SCI.
FT      Clinton G, Henner WD, Evans A;
FT      WPI; 2001-529934/58.
FT      New polypeptide, which binds to the extracellular domain of HER-2 for
FT      the treatment of hard tumors -
FT      Example 11; Page -: 61pp; English.
FT      The invention relates to novel HER-2 (herstatin-2) antagonist
FT      particularly a polypeptide that binds to the extracellular domain (ECD)
FT      of HER-2 at a site that is different from the binding site of humanised
FT      antibody, Herceptin, at an affinity of at least 10-8. The present
FT      invention is based upon the initial discovery of an alternative HER-2
FT      mRNA transcript with 274 bp insert of intron 8. The translation product
FT      of the alternative transcript is a truncated HER-2 protein designated
FT      p68HER-2 which lacks the transmembrane and intracellular domains of
FT      p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
FT      The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
FT      the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
FT      nucleic acids encoding these are useful to treat, diagnose and identify
FT      solid tumours. The present sequence is human p68HER-2 generic protein
FT      containing ECDIIIA variant sequence.
FT      Note: The present sequence is not shown in the specification but is
FT      derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
FT      sequence listing (AAE09181).
FT      XX
FT      SQ Sequence 419 AA:
FT      Query Match 95.4%; Score 2181; DB 22; Length 419;
FT      Best Local Similarity 95.9%; Pred. No. 7, 2e-169;
FT      Matches 402; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```

```

QY      181 LLLIDNRSRACHPCSPMKGRSGWSESSDCSLRTVACAGCARCKPLPTDCHEQC 240
QY      181 LLLIDNRSRACHPCSPMKGRSGWSESSDCSLRTVACAGCARCKPLPTDCHEQC 240
QY      241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNNDTFESMNPNGRRTTFGASCYTACP 300
QY      241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNNDTFESMNPNGRRTTFGASCYTACP 300
QY      301 YNYLSTDVGSCITVCPHLNHOEVAEDGTORCEKSKPCARGTSHLLPRPAAPVPLRMOP 360
QY      301 YNYLSTDVGSCITVCPHLNHOEVAEDGTORCEKSKPCARGTSHLLPRPAAPVPLRMOP 360
QY      361 GPAHPVLSFLRPSWDLVSAFYSLLAPLSPTSVPIPSVSGRGCPDDAHVAANLSRYEG 419
QY      361 GPAHPVLSFLRPSWDLVSAFYSLLAPLSPTSVPIPSVSGRGCPDDAHVAANLSRYEG 419
DB      361 XPAHPVLSFLRPSWDLVSAFYSLLAPLSPTSVPIPSVSGRGCPDDAHVAANLSRYEG 419
DB      361 XPAHPVLSFLRPSWDLVSAFYSLLAPLSPTSVPIPSVSGRGCPDDAHVAANLSRYEG 419
RESULT 2
AAE09212
ID      AAE09212 standard; Protein; 419 AA.
AC      AAE09212;
XX
XX      15-NOV-2001 (first entry)
XX
XX      Human p68HER-2 generic protein variant 10.
XX
XX      HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX      solid tumour; cancer; polymorphism; cytostatic; gene therapy;
XX      p68HER-2; ECDIIIA; variant.
XX
XX      Homo sapiens.
XX
XX      Key Location/Qualifiers
XX      FH 1..340
XX      FT Region /note="Identical to N-terminal region of p185HER-2"
XX      FT Domain 341..419
XX      FT /label= ECDIIIA-variant
XX      FT /note="Extracellular domain IIA variant"
XX      FT Misc-difference 124 /note="Represented as Agn in the parent sequence shown
XX      in the specification"
XX      FT Misc-difference 125 /note="Represented as Agn in the parent sequence shown
XX      in the specification"
XX      FT Misc-difference 342 /note="Represented as Agn in the parent sequence shown
XX      in the specification"
XX      FT Misc-difference 345 /label= Unknown
XX      FT Misc-difference 346 /label= Unknown
XX      FT Misc-difference 346 /label= Unknown
XX      FT Misc-difference 356 /label= Unknown
XX      FT Misc-difference 356 /label= Unknown
XX      FT Misc-difference 358 /label= Unknown
XX      FT Misc-difference 358 /label= Unknown
XX      FT Misc-difference 361 /label= Unknown
XX      FT Misc-difference 361 /label= Unknown
XX      FT Misc-difference 376 /label= Unknown
XX      FT Misc-difference 376 /label= Unknown
XX      FT Misc-difference 394 /label= Unknown
XX      FT Misc-difference 394 /label= Unknown
XX      FT Misc-difference 404 /label= Unknown
XX      FT Misc-difference 404 /label= Unknown
XX      FT Misc-difference 413 /label= Unknown
XX      FT /note="p68HER-2 generic sequence (AAE09181) Xaa
XX      substituted with Asn"
XX      WO200161356-A1.
XX      23-AUG-2001.
XX      16-FEB-2001; 2001WO-US05327.
XX

```

PR 16-FEB-2000; 2000US-0506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
XX
XX WPI: 2001-529934/58.
XX
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for
XX the treatment of hard tumors -
XX
XX Example 11; Page -: 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Hereceptin, at an affinity of at least 10⁷8. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 generic protein
XX containing ECDIIIA variant sequence.
XX Note: The present sequence is not shown in the specification but is
XX derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
XX sequence listing (AAE09181).
XX
XX SQ Sequence 419 AA:
XX
XX
XX Query Match 95.1%; Score 2176; DB 22; Length 419;
XX Best Local Similarity 95.7%; Pred. No. 1.8e-168;
XX Matches 401; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
XX
XX
XX 1 METALCRMGGLLALLPFAASTOVCTGDMKRLPASPETHLDMRLHYOGCQVVGNN 60
XX 1 METALCRMGGLLALLPFAASTOVCTGDKIRLPASETHLDMRLHYOGCQVVGNN 60
XX
XX 61 ELTYLPTNASLFLDIQIEVGQVYLAHNOVROVPLQRLRIYRGTLQFEDNVALAVLDNG 120
XX 61 ELTYLPTNASLFLDIQIEVGQVYLAHNOVROVPLQRLRIYRGTLQFEDNVALAVLDNG 120
XX
XX 121 DPLNNTPTVYGASPGGLRELQRLSLFEILKGVLIQNRNQLCYQDTILMKDIFHKNNOLA 180
XX 121 DPLNNTPTVYGASPGGLRELQRLSLFEILKGVLIQNRNQLCYQDTILMKDIFHKNNOLA 180
XX
XX 181 LTLIDNRSRACHPCSPMCKSGRCMGESSEDCOSLFRVCAGAGCARGKPLPTDCHEOC 240
XX 181 LTLIDNRSRACHPCSPMCKSGRCMGESSEDCOSLFRVCAGAGCARGKPLPTDCHEOC 240
XX
XX 241 AAGCTGPKHSDCLACLFHNSGICELCPALVYNTDTFESMPNPEGRYTFGASCVTACP 300
XX 241 AAGCTGPKHSDCLACLFHNSGICELCPALVYNTDTFESMPNPEGRYTFGASCVTACP 300
XX
XX 301 YNVLSTDVSGSCTLVCPHNOEYTAEDGTQRCCEKSPKARGHSLPRAAIVPPLRMQP 360
XX 301 YNVLSTDVSGSCTLVCPHNOEYTAEDGTQRCCEKSPKARGHSLPRAAIVPPLRMQP 360
XX
XX 361 GPAHPVLSFLRSPMDLVSAFYSLLPLAPLSPTSVPIPSVVGROPPDAVAVALSYEEG 419
XX 361 GPAHPVLSFLRSPMDLVSAFYSLLPLAPLSPTSVPIPSVVGROPPDAVAVALSYEEG 419
XX
XX
XX RESULT 3
XX AAE09205
XX ID AAE09205 standard; Protein: 419 AA.
XX AC AAE09205;
XX XX
XX 15-NOV-2001 (first entry)

XX
XX Human p68HER-2 generic protein variant 3.
XX
XX DE HER-2; herstatin; antagonist; extracellular domain; ECD; Hereceptin;
XX KW Solid tumour; cancer; polymorphism; cytostatic; gene therapy;
XX KW p68HER-2; ECDIIIA; variant.
XX
XX OS Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX Region 1..340
XX /note="Identical to N-terminal region of p185HER-2"
XX Domain 341..419
XX /label="ECDIIIA variant"
XX /note="Extracellular domain IIIA variant"
XX Misc-difference 124
XX /note="Represented as Agn in the parent sequence shown
XX in the specification"
XX Misc-difference 125
XX /note="Represented as Agn in the parent sequence shown
XX in the specification"
XX Misc-difference 342
XX /label="Unknown"
XX Misc-difference 345
XX /label="Unknown"
XX Misc-difference 346
XX /note="p68HER-2 generic sequence (AAE09181) Xaa
XX substituted with Leu"
XX Misc-difference 356
XX /label="Unknown"
XX Misc-difference 358
XX /label="Unknown"
XX Misc-difference 361
XX /label="Unknown"
XX Misc-difference 376
XX /label="Unknown"
XX Misc-difference 394
XX /label="Unknown"
XX Misc-difference 404
XX /label="Unknown"
XX Misc-difference 413
XX /label="Unknown"
XX W0200161356-A1.
XX
XX 23-AUG-2001.
XX PD
XX 16-FEB-2001; 2001WO-US05327.
XX PF
XX 16-FEB-2000; 2000US-0506079.
XX PR
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX PA
XX Clinton G, Henner WD, Evans A;
XX
XX WPI: 2001-529934/58.
XX
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for
XX the treatment of hard tumors -
XX
XX Example 11; Page -: 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Hereceptin, at an affinity of at least 10⁷8. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the

CC nucleic acids encoding these are useful to treat, diagnose and identify
 CC solid tumours. The present sequence is human p68HER-2 generic protein
 CC containing ECDIIIA variant sequence.
 CC Note: The present sequence is not shown in the specification but is
 CC derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
 CC sequence listing (AAE09181).

SO Sequence 419 AA:

Query Match 95.1%; Score 2174; DB 22; Length 419;
 Best Local Similarity 95.7%; Pred. No. 2.7e-168;
 Matches 401; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

OY 1 MELAALCRWGLLALLPPGAASVQCTGDMKRLRASPETHLDMIRHLYOGCQVVGNL 60
 DB 1 MELAALCRWGLLALLPPGAASVQCTGDMKRLRASPETHLDMIRHLYOGCQVVGNL 60
 OY 61 ELTYLPTNASLFLQDIOEVQGVYLAHNOVROPLORLRIYRGTLQFEDNYALAVLDNG 120
 DB 61 ELTYLPTNASLFLQDIOEVQGVYLAHNOVROPLORLRIYRGTLQFEDNYALAVLDNG 120
 OY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQRNPQLCYQDTILMKDIFHKNNOLA 180
 DB 121 DPLXHTPTVTGASPGGLRELQRLSLTEILKGVLIQRNPQLCYQDTILMKDIFHKNNOLA 180
 OY 181 LTLIDTNRSRACHPCSPMKSGSRMGESSPDQSLTRTVACAGSCARCKGPLPTDCCHQC 240
 DB 181 LTLIDTNRSRACHPCSPMKSGSRMGESSPDQSLTRTVACAGSCARCKGPLPTDCCHQC 240
 OY 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMPNPGRYTFGASCVTACP 300
 DB 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMPNPGRYTFGASCVTACP 300
 OY 301 YNTLSTDVSGCTVYCLHNOEYTAEDGTQCEKSKPCARGTSLPRPAAVVPVPLRMQP 360
 DB 301 YNTLSTDVSGCTVYCLHNOEYTAEDGTQCEKSKPCARGTSLPRPAAVVPVPLRMQP 360
 OY 361 GPAHPVLSFLRSPMDLVSAFYSLPLAPLSPTSVISPVSGRGDPDDAHAAVNLISRYEG 419
 DB 361 GPAHPVLSFLRSPMDLVSAFYSLPLAPLSPTSVISPVSGRGDPDDAHAAVNLISRYEG 419

RESULT 4
 AAE09207
 ID AAE09207 standard; Protein: 419 AA.

AC AAE09207;
 XX 15-NOV-2001 (first entry)
 DT Human p68HER-2 generic protein variant 5.
 DE
 XX
 KM HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
 KW solid tumor; cancer; polymorphism; cytostatic; gene therapy;
 KM p68HER-2; ECDIIIA; variant.
 XX
 OS Homo sapiens.
 XX

Key Location/Qualifiers
 FH 1..340
 FT Region /note= "Identical to N-terminal region of p185HER-2"
 FT Domain 341..419
 FT /label= ECDIIIA variant
 FT /note= "Extracellular domain IIIA variant"
 FT Misc-difference 124
 FT /note= "Represented as Agn in the parent sequence shown
 FT in the specification"
 FT Misc-difference 125
 FT /note= "Represented as Agn in the parent sequence shown
 FT in the specification"
 FT Misc-difference 342
 FT /label= unknown
 FT Misc-difference 345

FT /label= unknown
 FT Misc-difference 346
 FT /label= unknown
 FT Misc-difference 356
 FT /label= unknown
 FT Misc-difference 358
 FT /note= "p68HER-2 generic sequence (AAE09181) Xaa
 FT substituted with Leu"
 FT Misc-difference 361
 FT /label= unknown
 FT Misc-difference 376
 FT /label= unknown
 FT Misc-difference 394
 FT /label= unknown
 FT Misc-difference 404
 FT /label= unknown
 FT Misc-difference 413
 FT /label= unknown
 XX WO200161356-A1.
 XX
 PD 23-AUG-2001.
 XX
 PF 16-FEB-2001; 2001WO-US05327.
 XX
 PR 16-FEB-2000; 2000US-0506079.
 XX
 PA (UYOR-) UNIV OREGON HEALTH SCI.
 XX
 PI Clinton G, Henner WD, Evans A;
 DR WPI: 2001-529934/58.
 XX
 PT New polypeptide, which binds to the extracellular domain of HER-2 for
 the treatment of hard tumors .
 XX
 PS Example 11; Page -: 61pp; English.
 CC The invention relates to novel HER-2 (herstatin-2) antagonist
 CC particularly a polypeptide that binds to the extracellular domain (ECD)
 CC of HER-2 at a site that is different from the binding site of humanised
 CC antibody, Herceptin, at an affinity of at least 10⁻⁸. The present
 CC invention is based upon the initial discovery of an alternative HER-2
 CC mRNA transcript with 274 bp insert of intron 8. The translation product
 CC of the alternative transcript is a truncated HER-2 protein designated
 CC p68HER-2 which lacks the transmembrane and intracellular domains of
 CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
 CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
 CC nucleic acids encoding these are useful to treat, diagnose and identify
 CC solid tumours. The present sequence is human p68HER-2 generic protein
 CC containing ECDIIIA variant sequence.
 CC Note: The present sequence is not shown in the specification but is
 CC derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
 CC sequence listing (AAE09181).
 XX
 SO Sequence 419 AA:
 Query Match 95.0%; Score 2172; DB 22; Length 419;
 Best Local Similarity 95.5%; Pred. No. 3.9e-168;
 Matches 400; Conservative 1; Mismatches 18; Indels 0; Gaps 0;


```
QY 181 LTLIDNRSRACHPCSPCKGSRGWGESSSDCQSLTRTVACAGGACRCKGPLPTDCCHEQC 240
    |||||||
Db 181 LTLIDNRSRACHPCSPCKGSRGWGESSSDCQSLTRTVACAGGACRCKGPLPTDCCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHFNHSGICELHCPALVYNTDFESMPNPEGRTFGASCVTACP 300
    |||||||
Db 241 AAGCTGPKHSDCLACLFHFNHSGICELHCPALVYNTDFESMPNPEGRTFGASCVTACP 300
QY 301 YNLTSDVGSCTLVCPHLNQEVTAEADGTORCEKSRPCAGTSHLLRPAAVVPLRMOP 360
    |||||||
Db 301 YNLTSDVGSCTLVCPHLNQEVTAEADGTORCEKSRPCAGTSHLLRPAAVVPLRMOP 360
QY 361 GPAHPVLSFLRPSMDLVSAFYSLPLAPLSPTSPISPVSGRGPDPDAHVAVNLSTREG 419
    |||||||
Db 361 GPAHPVLSFLRPSMDLVSAFYSLPLAPLSPTSPISPVSGRGPDPDAHVAVNLSTREG 419

RESULT 5
AAE09209 ID AAE09209 standard; Protein; 419 AA.
XX AC AAE09209;
XX DT 15-NOV-2001 (first entry)
XX DE Human p68HER-2 generic protein variant 7.
XX KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;
XX KW p68HER-2; ECDIIIA; variant.
XX OS Homo sapiens.
XX PH Key Location/Qualifiers
    FT 1..340
    FT /note= "Identical to N-terminal region of p185HER-2"
    FT Domain 341..419
    FT /label= "ECDIIIA; variant"
    FT /note= "Extracellular domain IIIA variant"
    FT Misc-difference 124
    FT /note= "Represented as Agn in the parent sequence shown
    FT in the specification"
    FT Misc-difference 125
    FT /note= "Represented as Agn in the parent sequence shown
    FT in the specification"
    FT Misc-difference 342
    FT /label= Unknown
    FT Misc-difference 345
    FT /label= Unknown
    FT Misc-difference 346
    FT /label= Unknown
    FT Misc-difference 356
    FT /label= Unknown
    FT Misc-difference 358
    FT /label= Unknown
    FT Misc-difference 361
    FT /label= Unknown
    FT Misc-difference 376
    FT /note= "p68HER-2 generic sequence (AAE09181) Xaa
    FT substituted with Ile"
    FT Misc-difference 394
    FT /label= Unknown
    FT Misc-difference 404
    FT /label= Unknown
    FT Misc-difference 413
    FT /label= Unknown
    XX WO200161356-A1.
    XX 23-AUG-2001.
    XX 16-FEB-2001; 2001WO-US05327.
    PF
```

```
XX 16-FEB-2000; 2000US-0506079.
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX Clinton G, Henner WD, Evans A;
XX WPI; 2001-529934/58.
XX New polypeptide, which binds to the extracellular domain of HER-2 for
XX the treatment of hard tumors .
XX Example 11; Page -: 61pp; English.
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 10-8. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 generic protein
XX containing ECDIIIA variant sequence.
XX Note: The present sequence is not shown in the specification but is
XX derived from p68HER-2 generic sequence (SBO ID NO:2) shown in the
XX sequence listing (AAE09181).
XX SQ Sequence 419 AA:
```

```
Query Match 95.0%; Score 2172; DB 22; Length 419;
Best Local Similarity 95.5%; Pred. No. 3,9e-168;
Matches 400; Conservative 1; Mismatches 18; Indels 0; Gaps 0;
```

```
QY 1 MEALALCRWGLLALLPRGASTOVCTGTDMLRLASPEITHIDMLRHLYGCGVVGNI 60
    |||||||
Db 1 MEALALCRWGLLALLPRGASTOVCTGTDCKLRLASPEITHIDMLRHLYGCGVVGNI 60
QY 61 ELTYLPTNASTSLFQDIQEVGYVLLAHNOVROVPLORLRIVGTOLFEDNVLAFLVNG 120
    |||||||
Db 61 ELTYLPTNASTSLFQDIQEVGYVLLAHNOVROVPLQRLRIYVGTOLFEDNVLAFLVNG 120
QY 121 DPLNMTPTVTGASPGGLREIQRLSLEILKGVLLIQRNPLQCYQDTILMKDIFHKNNOLA 180
    |||||||
Db 121 DPLXHTPTVTGASPGGLREIQRLSLEILKGVLLIQRNPLQCYQDTILMKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPCKGSRGWGESSSDCQSLTRTVACAGGACRCKGPLPTDCCHEQC 240
    |||||||
Db 181 LTLIDNRSRACHPCSPCKGSRGWGESSSDCQSLTRTVACAGGACRCKGPLPTDCCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHFNHSGICELHCPALVYNTDFESMPNPEGRTFGASCVTACP 300
    |||||||
Db 241 AAGCTGPKHSDCLACLFHFNHSGICELHCPALVYNTDFESMPNPEGRTFGASCVTACP 300
QY 301 YNLTSDVGSCTLVCPHLNQEVTAEADGTORCEKSRPCAGTSHLLRPAAVVPLRMOP 360
    |||||||
Db 301 YNLTSDVGSCTLVCPHLNQEVTAEADGTORCEKSRPCAGTSHLLRPAAVVPLRMOP 360
QY 361 GPAHPVLSFLRPSMDLVSAFYSLPLAPLSPTSPISPVSGRGPDPDAHVAVNLSTREG 419
    |||||||
Db 361 GPAHPVLSFLRPSMDLVSAFYSLPLAPLSPTSPISPVSGRGPDPDAHVAVNLSTREG 419

RESULT 6
AAE20348 ID AAE20348 standard; Protein; 419 AA.
XX AC AAE20348;
XX
```

DT	18-JUN-2002	(first entry)
XX	Human truncated HER2 protein.	
DE	Human truncated HER2 protein.	
XX	Human: tumour; endothelial growth factor receptor; EGFR; cytosolic; herstatin; HER-2 receptor tyrosine kinase; squamous cell carcinoma; lung; colon; glial cell tumour; cell growth.	
KW	Human: tumour; endothelial growth factor receptor; EGFR; cytosolic; herstatin; HER-2 receptor tyrosine kinase; squamous cell carcinoma; lung; colon; glial cell tumour; cell growth.	
RW	Human: tumour; endothelial growth factor receptor; EGFR; cytosolic; herstatin; HER-2 receptor tyrosine kinase; squamous cell carcinoma; lung; colon; glial cell tumour; cell growth.	
XX	Human: tumour; endothelial growth factor receptor; EGFR; cytosolic; herstatin; HER-2 receptor tyrosine kinase; squamous cell carcinoma; lung; colon; glial cell tumour; cell growth.	
OS	Homo sapiens.	
XX	Homo sapiens.	
Key	Location/Qualifiers	
FH	Misc-difference 342	
FT	/label= Thr, Ser	
FT	Misc-difference 345	
FT	/label= Leu, Pro	
FT	Misc-difference 346	
FT	/label= Pro, Leu	
FT	Misc-difference 356	
FT	/label= Leu, Gln	
FT	Misc-difference 358	
FT	/label= Met, Leu	
FT	Misc-difference 361	
FT	/label= Gly, Asp, Ala, Val	
FT	Misc-difference 376	
FT	/label= Leu, Ile	
FT	Misc-difference 394	
FT	/label= Pro, Arg	
FT	Misc-difference 404	
FT	/label= Pro, Leu	
FT	Misc-difference 413	
FT	/label= Asp, Asn	
XX	WO200214470-A2.	
PN	WO200214470-A2.	
XX	21-FEB-2002.	
PD	21-FEB-2002.	
XX	14-AUG-2001; 2001WO-US25502.	
PF	14-AUG-2001; 2001WO-US25502.	
XX	14-AUG-2000; 2000US-0638834.	
PR	14-AUG-2000; 2000US-0638834.	
XX	(UYOR-) UNIV OREGON HEALTH SCI.	
XX	(UYOR-) UNIV OREGON HEALTH SCI.	
PA	(UYOR-) UNIV OREGON HEALTH SCI.	
XX	Clinton GM.	
PI	Clinton GM.	
XX	WPI; 2002-269185/31.	
DR	WPI; 2002-269185/31.	
XX	Treating solid tumor characterized by expression of endothelial growth factor receptor, involves administering recombinant herstatin that binds to extracellular domain of the endothelial growth factor receptor	
PT	Treating solid tumor characterized by expression of endothelial growth factor receptor, involves administering recombinant herstatin that binds to extracellular domain of the endothelial growth factor receptor	
PT	Treating solid tumor characterized by expression of endothelial growth factor receptor, involves administering recombinant herstatin that binds to extracellular domain of the endothelial growth factor receptor	
PT	Treating solid tumor characterized by expression of endothelial growth factor receptor, involves administering recombinant herstatin that binds to extracellular domain of the endothelial growth factor receptor	
XX	Claim 1; Page 78-80; 82pp; English.	
PS	Claim 1; Page 78-80; 82pp; English.	
XX	The present invention relates to a method for treating a solid tumour characterised by endothelial growth factor receptor (EGFR) expression. The method involves administering an agent that binds to an extracellular domain (ECD) of EGFR. The invention also relates to a naturally occurring inhibitor of HER-2 receptor tyrosine kinase called herstatin. The co-expression of herstatin with p185HER2 causes a striking reduction in cell growth that corresponds with suppression of p185 autophosphorylation. The method or a pharmaceutical composition is useful for treating a solid tumour (selected from squamous cell carcinoma, lung carcinoma, colon carcinoma and glial cell tumour) characterised by EGFR expression. The present sequence is human truncated HER2 protein that lacks transmembrane and intracellular domains.	
CC	The present invention relates to a method for treating a solid tumour characterised by endothelial growth factor receptor (EGFR) expression. The method involves administering an agent that binds to an extracellular domain (ECD) of EGFR. The invention also relates to a naturally occurring inhibitor of HER-2 receptor tyrosine kinase called herstatin. The co-expression of herstatin with p185HER2 causes a striking reduction in cell growth that corresponds with suppression of p185 autophosphorylation. The method or a pharmaceutical composition is useful for treating a solid tumour (selected from squamous cell carcinoma, lung carcinoma, colon carcinoma and glial cell tumour) characterised by EGFR expression. The present sequence is human truncated HER2 protein that lacks transmembrane and intracellular domains.	
CC	The present invention relates to a method for treating a solid tumour characterised by endothelial growth factor receptor (EGFR) expression. The method involves administering an agent that binds to an extracellular domain (ECD) of EGFR. The invention also relates to a naturally occurring inhibitor of HER-2 receptor tyrosine kinase called herstatin. The co-expression of herstatin with p185HER2 causes a striking reduction in cell growth that corresponds with suppression of p185 autophosphorylation. The method or a pharmaceutical composition is useful for treating a solid tumour (selected from squamous cell carcinoma, lung carcinoma, colon carcinoma and glial cell tumour) characterised by EGFR expression. The present sequence is human truncated HER2 protein that lacks transmembrane and intracellular domains.	
CC	The present invention relates to a method for treating a solid tumour characterised by endothelial growth factor receptor (EGFR) expression. The method involves administering an agent that binds to an extracellular domain (ECD) of EGFR. The invention also relates to a naturally occurring inhibitor of HER-2 receptor tyrosine kinase called herstatin. The co-expression of herstatin with p185HER2 causes a striking reduction in cell growth that corresponds with suppression of p185 autophosphorylation. The method or a pharmaceutical composition is useful for treating a solid tumour (selected from squamous cell carcinoma, lung carcinoma, colon carcinoma and glial cell tumour) characterised by EGFR expression. The present sequence is human truncated HER2 protein that lacks transmembrane and intracellular domains.	
CC	The present invention relates to a method for treating a solid tumour characterised by endothelial growth factor receptor (EGFR) expression. The method involves administering an agent that binds to an extracellular domain (ECD) of EGFR. The invention also relates to a naturally occurring inhibitor of HER-2 receptor tyrosine kinase called herstatin. The co-expression of herstatin with p185HER2 causes a striking reduction in cell growth that corresponds with suppression of p185 autophosphorylation. The method or a pharmaceutical composition is useful for treating a solid tumour (selected from squamous cell carcinoma, lung carcinoma, colon carcinoma and glial cell tumour) characterised by EGFR expression. The present sequence is human truncated HER2 protein that lacks transmembrane and intracellular domains.	
XX	Sequence 419 AA:	
SO	Sequence 419 AA:	
Query Match	94.9%; Score 2171; DB 23; Length 419;	
Best Local Similarity	95.5%; Pred. No. 4,7e-168;	
Matches 400: Conservative	0; Mismatches 19; Indels 0; Gaps 0	

Db	1	MEALACRMGLLALLPRGCASTQVCTGDKLRLPASPETHIDMLRHLVQSGOVVQGNL	60
Qy	61	ELTYLPTNASLSFIODIOEVQGYVLIAHNOVRLQRLRIYRGTOLEEDNYALAVLNG	120
Db	61	ELTYLPTNASLSFIODIOEVQGYVLCAHNOVRLQRLRIYRGTOLEEDNYALAVLNG	120
Qy	121	DPLNNTVTYVTSAGSPGGLAEFLDLRLSTELLKGVLLIQRNPLCYQDTILMKDIFHKNNOLA	180
Db	121	DPLRRTVTYVTSAGSPGGLAEFLDLRLSTELLKGVLLIQRNPLCYQDTILMKDIFHKNNOLA	180
Qy	181	LTLIDTNSRACHPCSPCKCKSGRCGSESEDDQSLTRFYCAGCARGKPLPTDCCHEC	240
Db	181	LTLIDTNSRACHPCSPCKCKSGRCGSESEDDQSLTRFYCAGCARGKPLPTDCCHEC	240
Qy	241	AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNEEGRYTFGASCVTACP	300
Db	241	AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNEEGRYTFGASCVTACP	300
Qy	301	YNYLSTDVSGCTLYCPRLHNOEYTLADGTQRCCKSPCARGTHSLLPRAAPVPLRMQP	360
Db	301	YNYLSTDVSGCTLYCPRLHNOEYTLADGTQRCCKSPCARGTHSLLPRAAPVPLRMQP	360
Qy	361	GPAPHVLSFLRPSMWLVSAFYSLPLAPSPTSVPSPSPVSGRPPDMAVAVNLSRYEG	419
Db	361	XPAHPLYSLRPSMWLVSAFYSLPLAPLPDPTSVXISPVSGRGXDPDAHVAVXLSRYEG	419
RESULT 7			
AAE09203	ID	AAE09203 standard; Protein; 419 AA.	
XX	AC	AAE09203;	
XX	DT	15-NOV-2001 (first entry)	
DE	XX	Human p68HER-2 generic protein variant 1.	
XX	KW	HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;	
KW	KW	solid tumour; cancer; polymorphism; cytostatic; gene therapy;	
KW	KW	p68HER-2; ECDIIia; variant.	
OS	XX	Homo sapiens.	
XX	XX		
FT	FT	Key	Location/Qualifiers
FT	FT	Region	1..340
FT	FT	Domain	/note= "Identical to N-terminal region of p185HER-2"
FT	FT		341..419
FT	FT		/label= ECDIIia_variant
FT	FT	Misc-difference	/note= "Extracellular domain IIIa variant"
FT	FT		124
FT	FT		/note= "Represented as Agn in the parent sequence shown
FT	FT		in the specification"
FT	FT	Misc-difference	125
FT	FT		/note= "Represented as Agn in the parent sequence shown
FT	FT		in the specification"
FT	FT	Misc-difference	342
FT	FT		/note= "p68HER-2 generic sequence (AAE09181) Xaa
FT	FT		substituted with Ser"
FT	FT	Misc-difference	345
FT	FT		/label= Unknown
FT	FT	Misc-difference	346
FT	FT		/label= Unknown
FT	FT	Misc-difference	356
FT	FT		/label= Unknown
FT	FT	Misc-difference	358
FT	FT		/label= Unknown
FT	FT	Misc-difference	361
FT	FT		/label= Unknown
FT	FT	Misc-difference	376
FT	FT		/label= Unknown
FT	FT	Misc-difference	394
FT	FT		/label= Unknown

```

FT Misc-difference 404
FT /label= Unknown
FT Misc-difference 413
FT /label= Unknown
XX
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US05327.
XX
XX 16-FEB-2000; 2000US-0506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
XX
XX WPI; 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for
XX the treatment of hard tumors -
XX
XX Example 11, Page -: 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Hereceptin, at an affinity of at least 10-8. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 generic protein
XX containing ECDIIIA variant sequence.
XX Note: The present sequence is not shown in the specification but is
XX derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
XX sequence listing (AAE09181).
XX
XX Sequence 419 AA:
XX
XX Query Match 94.9%; Score 2170; DB 22; Length 419;
XX Best Local Similarity 95.5%; Pred No. 5, 7e-168;
XX Matches 400; Conservative 1; Mismatches 18; Indels 0; Gaps 0;

```

```

DB 361 XPAHPVLSFLRPSMDVSAFYSLPLAPLPTSVXISPVSGRGXDPDAHVAVXLSRYEG 419
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
RESULT 8
ID AAE09181 standard; Protein; 419 AA.
XX
XX AAE09181;
XX
XX 15-NOV-2001 (first entry)
XX
XX Human p68HER-2 generic sequence #1.
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Hereceptin;
XX solid tumour; cancer; polymorphism; cytostatic; gene therapy;
XX p68HER-2; ECDIIIA.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH 1..340
XX FT /note= "Identical to N-terminal region of p185HER-2"
XX FT Domain 341..419
XX FT /label= ECDIIIA
XX FT /note= "Extracellular domain IIIA"
XX
XX Misc-difference 124 /note= "Represented as Agn in the sequence shown in
XX the specification"
XX
XX Misc-difference 125 /note= "Represented as Agn in the sequence shown in
XX the specification"
XX
XX Misc-difference 342 /label= Unknown
XX FT /note= "Encoded by WCC"
XX
XX Misc-difference 345 /label= Unknown
XX FT /note= "Encoded by CYG"
XX
XX Misc-difference 346 /label= Unknown
XX FT /note= "Encoded by CYC"
XX
XX Misc-difference 356 /label= Unknown
XX FT /note= "Encoded by CYC"
XX
XX Misc-difference 358 /label= Unknown
XX FT /note= "Encoded by CMC"
XX
XX Misc-difference 361 /label= Unknown
XX FT /note= "Encoded by GNC"
XX
XX Misc-difference 376 /label= Unknown
XX FT /note= "Encoded by WTA"
XX
XX Misc-difference 389 /note= "Encoded by AGC"
XX
XX Misc-difference 394 /label= Unknown
XX FT /note= "Encoded by CST"
XX
XX Misc-difference 404 /label= Unknown
XX FT /note= "Encoded by CYG"
XX
XX Misc-difference 413 /label= Unknown
XX FT /note= "Encoded by SAC"
XX
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US05327.
XX
XX 16-FEB-2000; 2000US-0506079.
XX
XX
XX

```

PA (UYOR-) UNIV OREGON HEALTH SCI.
 XX Clinton G, Henner WD, Evans A;
 PI
 XX
 DR WPI: 2001-529934/58.
 DR N-PSDB; AAD15844.
 XX
 PT New polypeptide, which binds to the extracellular domain of HER-2 for
 XX the treatment of hard tumors -
 PS
 XX Claim 8; Page 53-54; 61pp; English.
 CC The invention relates to novel HER-2 (herstatin-2) antagonist
 CC particularly a polypeptide that binds to the extracellular domain (ECD)
 CC of HER-2 at a site that is different from the binding site of humanised
 CC antibody, Herceptin, at an affinity of at least 10⁻⁸. The present
 CC invention is based upon the initial discovery of an alternative HER-2
 CC mRNA transcript with 274 bp insert of intron 8. The translation product
 CC of the alternative transcript is a truncated HER-2 protein designated
 CC p68HER-2 which lacks the transmembrane and intracellular domains of
 CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
 CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
 CC nucleic acids encoding these are useful to treat, diagnose and identify
 CC solid tumours. The present sequence is human p68HER-2 protein
 CC containing ECDIIIA generic sequence.
 XX
 SO Sequence 419 AA;
 Query Match 94.8%; Score 2169; DB 22; Length 419;
 Best Local Similarity 95.5%; Pred. No. 6,8e-168;
 Matches 400; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
 Oy 1 MELAALCRMGLLALLPPGAATOVCTGTDMKRLPASPETHDMLRHLVQGCVOVQGNL 60
 Db 1 MELAALCRMGLLALLPPGAATOVCTGTDCIKRLPASPETHDMLRHLVQGCVOVQGNL 60
 Oy 61 ELTYLPTNASLSFLQDIQVQGYVLIANQVRQVPLQRLRIVRGTLFEDNVALAVLNG 120
 Db 61 ELTYLPTNASLSFLQDIQVQGYVLIANQVRQVPLQRLRIVRGTLFEDNVALAVLNG 120
 Oy 121 DPLNNTPTVYGASPGGLRELQLSLEILKGVLIQBNPOLCYQDTILMKDITHKNNOLA 180
 Db 121 DPLNNTPTVYGASPGGLRELQLSLEILKGVLIQBNPOLCYQDTILMKDITHKNNOLA 180
 Oy 121 DPLNNTPTVYGASPGGLRELQLSLEILKGVLIQBNPOLCYQDTILMKDITHKNNOLA 180
 Db 121 DPLNNTPTVYGASPGGLRELQLSLEILKGVLIQBNPOLCYQDTILMKDITHKNNOLA 180
 Oy 181 LTLIDNRSRACPRCSGRCWGESSESDCSLTRTVAGGCARCKGFLPTDCCHEOC 240
 Db 181 LTLIDNRSRACPRCSGRCWGESSESDCSLTRTVAGGCARCKGFLPTDCCHEOC 240
 Oy 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCYTACP 300
 Db 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCYTACP 300
 Oy 301 YNLTLDVSGSCTVCPHLNQEYAEQGTQRCCKSPCARGTSLRPRAAVVPLRMOP 360
 Db 301 YNLTLDVSGSCTVCPHLNQEYAEQGTQRCCKSPCARGTSLRPRAAVVPLRMOP 360
 Oy 361 GPAHPVLSFLRPSMDLVSAFYSLPLAPLSTVPSVSGRGGDDPAHVAVLSRYEG 419
 Db 361 XPAHPVLSFLRPSMDLVSAFYSLPLAPLSTVPSVSGRGGDDPAHVAVLSRYEG 419
 RESULT 9
 AAE09208
 ID AAE09208 standard; Protein; 419 AA.
 XX
 AC AAE09208;
 XX
 DT 15-NOV-2001 (first entry)
 XX
 DE Human p68HER-2 generic protein variant 6.
 KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;

KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;
 KW p68HER-2; ECDIIIA; variant.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 1..340
 FT Domain /note= "Identical to N-terminal region of p185HER-2"
 FT /label= "ECDIIIA_variant"
 FT /note= "Extracellular domain IIIA variant"
 FT Misc-difference 124
 FT /note= "Represented as Agn in the parent sequence shown
 FT in the specification"
 FT Misc-difference 125
 FT /note= "Represented as Agn in the parent sequence shown
 FT in the specification"
 FT Misc-difference 342
 FT /label= "Unknown"
 FT Misc-difference 345
 FT /label= "Unknown"
 FT Misc-difference 346
 FT /label= "Unknown"
 FT Misc-difference 356
 FT /label= "Unknown"
 FT Misc-difference 358
 FT /label= "Unknown"
 FT Misc-difference 361
 FT /label= "Unknown"
 FT Misc-difference 376
 FT /label= "Asp, Ala, Val"
 FT Misc-difference 394
 FT /label= "Unknown"
 FT Misc-difference 404
 FT /label= "Unknown"
 FT Misc-difference 413
 FT /label= "Unknown"
 PN WO200161356-A1.
 XX
 PD 23-AUG-2001.
 XX
 XX 16-FEB-2001; 2001WO-US05327.
 XX
 XX 16-FEB-2000; 2000US-0506079.
 XX
 PA (UYOR-) UNIV OREGON HEALTH SCI.
 PI Clinton G, Henner WD, Evans A;
 XX
 PI WPI: 2001-529934/58.
 XX
 PT New polypeptide, which binds to the extracellular domain of HER-2 for
 XX the treatment of hard tumors -
 PS
 XX Example 11; Page -: 61pp; English.
 CC The invention relates to novel HER-2 (herstatin-2) antagonist
 CC particularly a polypeptide that binds to the extracellular domain (ECD)
 CC of HER-2 at a site that is different from the binding site of humanised
 CC antibody, Herceptin, at an affinity of at least 10⁻⁸. The present
 CC invention is based upon the initial discovery of an alternative HER-2
 CC mRNA transcript with 274 bp insert of intron 8. The translation product
 CC of the alternative transcript is a truncated HER-2 protein designated
 CC p68HER-2 which lacks the transmembrane and intracellular domains of
 CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
 CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
 CC nucleic acids encoding these are useful to treat, diagnose and identify
 CC solid tumours. The present sequence is human p68HER-2 generic protein
 CC containing ECDIIIA variant sequence.
 CC Note: The present sequence is not shown in the specification but is
 CC derived from p68HER-2 generic sequence (SEQ ID NO.2) shown in the

```

CC sequence listing (AAE09181).
XX
SQ Sequence 419 AA:
Query Match 94.8%; Score 2169; DB 22; Length 419;
Best Local Similarity 95.5%; Pred. No. 6.8e-168;
Matches 400; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

OY 1 MELAALCRWGLLALLPPCAASTOVCTGDMKRLPASPETHLDMRLHYOGCQVQGNL 60
DB 1 MELAALCRWGLLALLPPCAASTOVCTGDMKRLPASPETHLDMRLHYOGCQVQGNL 60
OY 61 ELTYLPTNASLSFLDIOEVQGVLLAHNOVQVPLQRLRIYRGTOLEFDNVALAVLDNG 120
DB 61 ELTYLPTNASLSFLDIOEVQGVLLAHNOVQVPLQRLRIYRGTOLEFDNVALAVLDNG 120
OY 121 DPLNNTTPTVTGASPGGLRELOLRSLTEILKGVLIORNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTTPTVTGASPGGLRELOLRSLTEILKGVLIORNPOLCYODTILMKDIFHKNNOLA 180
OY 181 LRLIDNRSRACHPCSPCKGSRWGESESDCOSLTRVACAGCARKGKPLPTDCCHEC 240
DB 181 LRLIDNRSRACHPCSPCKGSRWGESESDCOSLTRVACAGCARKGKPLPTDCCHEC 240
OY 241 AAGCTGPKHSDCLAHFNHSGICELCPALVYNTDFESMPNPGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLAHFNHSGICELCPALVYNTDFESMPNPGRYTFGASCVTACP 300
OY 301 YNLTSDVSCSLVCPPLHNOEYTAEDGTORCEKSKPCARGTHSLLPRAAPVPLRMOP 360
DB 301 YNLTSDVSCSLVCPPLHNOEYTAEDGTORCEKSKPCARGTHSLLPRAAPVPLRMOP 360
OY 361 GRAHPVLSFLRSPMDVSAFYSPLAPLSPTSVPISPVGREGPDDAHVAVALSYEG 419
DB 361 GRAHPVLSFLRSPMDVSAFYSPLAPLSPTSVPISPVGREGPDDAHVAVALSYEG 419

RESULT 10
AAE09210
ID AAE09210 standard; Protein: 419 AA.
XX
AC AAE09210;
XX
DT 15-NOV-2001 (first entry)
XX
DE Human p68HER-2 generic protein variant 8.
XX
KW HER-2; herstatin; antagonist; extracellular domain; ECD; Hereceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;
KW p68HER-2; ECDIIIA; variant.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH Region 1..340
FT /note= "Identical to N-terminal region of p185HER-2"
FT Domain 341..419
FT /label= "ECDIIIA_variant"
FT /note= "Extracellular domain IIA variant"
FT Misc-difference 124
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Misc-difference 125
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Misc-difference 342
FT /label= Unknown
FT Misc-difference 345
FT /label= Unknown
FT Misc-difference 346
FT /label= Unknown
FT Misc-difference 356
FT /label= Unknown

```

```

FT Misc-difference 358
FT /label= Unknown
FT Misc-difference 361
FT /label= Unknown
FT Misc-difference 376
FT /label= Unknown
FT Misc-difference 394
FT /note= "p68HER-2 generic sequence (AAE09181) Xaa
FT substituted with Arg"
FT Misc-difference 404
FT /label= Unknown
FT Misc-difference 413
FT /label= Unknown
PN WO200161356-A1.
PD 23-AUG-2001.
PF 16-FEB-2001; 2001WO-US05327.
PR 16-FEB-2000; 2000US-0506079.
PA (UYOR-) UNIV OREGON HEALTH SCI.
XX Clinton G, Henner WD, Evans A;
XX WPI; 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for
XX the treatment of hard tumors -
XX
XX Example 11: Page -: 61pp: English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Hereceptin, at an affinity of at least 10-8. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 generic protein
XX containing ECDIIIA variant sequence.
XX Note: The present sequence is not shown in the specification but is
XX derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
XX sequence listing (AAE09181).
XX
SQ Sequence 419 AA:
Query Match 94.8%; Score 2169; DB 22; Length 419;
Best Local Similarity 95.5%; Pred. No. 6.8e-168;
Matches 400; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

OY 1 MELAALCRWGLLALLPPCAASTOVCTGDMKRLPASPETHLDMRLHYOGCQVQGNL 60
DB 1 MELAALCRWGLLALLPPCAASTOVCTGDMKRLPASPETHLDMRLHYOGCQVQGNL 60
OY 61 ELTYLPTNASLSFLDIOEVQGVLLAHNOVQVPLQRLRIYRGTOLEFDNVALAVLDNG 120
DB 61 ELTYLPTNASLSFLDIOEVQGVLLAHNOVQVPLQRLRIYRGTOLEFDNVALAVLDNG 120
OY 121 DPLNNTTPTVTGASPGGLRELOLRSLTEILKGVLIORNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTTPTVTGASPGGLRELOLRSLTEILKGVLIORNPOLCYODTILMKDIFHKNNOLA 180
OY 181 LRLIDNRSRACHPCSPCKGSRWGESESDCOSLTRVACAGCARKGKPLPTDCCHEC 240
DB 181 LRLIDNRSRACHPCSPCKGSRWGESESDCOSLTRVACAGCARKGKPLPTDCCHEC 240

```

```
OY 241 AAGCTGPKHSDCLACLFHFNHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACP 300
    |||
DB 241 AAGCTGPKHSDCLACLFHFNHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACP 300
OY 301 YNLTSDVSGCTLVCPRLHNOEVTAEEDGTORCEKSPCARGTHSLPRPAAVVPRLMOP 360
    |||
DB 301 YNLTSDVSGCTLVCPRLHNOEVTAEEDGTORCEKSPCARGTHSLPRPAAVVPRLMOP 360
OY 361 GPAHPVLSFLRPSMDVSAFYSLPLAPLSPTSPVSPVSGRPPDDAHVAVNLSTYEG 419
    |||
DB 361 GPAHPVLSFLRPSMDVSAFYSLPLAPLSPTSPVSPVSGRPPDDAHVAVNLSTYEG 419

RESULT 11
AAE09206
ID AAE09206 standard: Protein: 419 AA.
XX
AC AAE09206;
XX
DT 15-NOV-2001 (first entry)
XX
DE Human p68HER-2 generic protein variant 4.
XX
KM HER-2; herstatin; antagonist; extracellular domain; ECD; Hereceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;
KM p68HER-2; ECDIIIA; variant.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 1..340
FT /note= "Identical to N-terminal region of p185HER-2"
FT Domain 341..419
FT /label= ECDIIIA variant
FT /note= "Extracellular domain IIIa variant"
FT Misc-difference 124
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Misc-difference 342
FT /label= unknown
FT Misc-difference 345
FT /label= unknown
FT Misc-difference 346
FT /label= unknown
FT Misc-difference 356
FT /note= "p68HER-2 generic sequence (AAE09181) Xaa
FT substituted with Gln"
FT Misc-difference 358
FT /label= unknown
FT Misc-difference 361
FT /label= unknown
FT Misc-difference 376
FT /label= unknown
FT Misc-difference 394
FT /label= unknown
FT Misc-difference 404
FT /label= unknown
FT Misc-difference 413
FT /label= unknown

WO200161356-A1.
PD 23-AUG-2001.
XX
PF 16-FEB-2001; 2001WO-0505327.
XX
PR 16-FEB-2000; 2000US-0506079.
XX
PA (UOR-) UNIV OREGON HEALTH SCI.
XX
```

```
PI Clinton G, Henner WD, Evans A;
XX
DR WPI: 2001-529934/58.
XX
PT New polypeptide, which binds to the extracellular domain of HER-2 for
PS the treatment of hard tumors.
XX
PS Example 11: Page -: 61pp: English.
XX
CC The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Hereceptin, at an affinity of at least 10-8. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is human p68HER-2 generic protein
CC containing ECDIIIA variant sequence.
CC Note: The present sequence is not shown in the specification but is
CC derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
CC sequence listing (AAE09181).
XX
SQ Sequence 419 AA:
XX
Query Match 94.8%; Score 2168; DB 22: Length 419;
Best Local Similarity 95.5%; Pred. No. 8,2e-168;
Matches 400; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
OY 1 MELAALCRWGLLLALLPRGAASNOVCTGDMKRLRPASPTFLDMIRHLXGCGVVGML 60
DB 1 MELAALCRWGLLLALLPRGAASNOVCTGDMKRLRPASPTFLDMIRHLXGCGVVGML 60
OY 61 ELTYLPTNLSFLDIDQEVQGVYLAHNOVQVPLQRLRIYRGQLFEDNALAVLDNG 120
DB 61 ELTYLPTNLSFLDIDQEVQGVYLAHNOVQVPLQRLRIYRGQLFEDNALAVLDNG 120
OY 121 DPLNNTPTVTGASPGGLRELRLSLTEILKGVLIORNPOLCYODTILMKDIFHKNOLA 180
DB 121 DPLNNTPTVTGASPGGLRELRLSLTEILKGVLIORNPOLCYODTILMKDIFHKNOLA 180
OY 121 DPLNNTPTVTGASPGGLRELRLSLTEILKGVLIORNPOLCYODTILMKDIFHKNOLA 180
DB 121 DPLNNTPTVTGASPGGLRELRLSLTEILKGVLIORNPOLCYODTILMKDIFHKNOLA 180
OY 181 LTLIDTNRSRACHPCSPMKSGSKWGESESDOSLRTYACAGCARCKGPLPTDCHEOC 240
DB 181 LTLIDTNRSRACHPCSPMKSGSKWGESESDOSLRTYACAGCARCKGPLPTDCHEOC 240
OY 181 LTLIDTNRSRACHPCSPMKSGSKWGESESDOSLRTYACAGCARCKGPLPTDCHEOC 240
DB 181 LTLIDTNRSRACHPCSPMKSGSKWGESESDOSLRTYACAGCARCKGPLPTDCHEOC 240
OY 241 AAGCTGPKHSDCLACLFHFNHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHFNHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACP 300
OY 301 YNLTSDVSGCTLVCPRLHNOEVTAEEDGTORCEKSPCARGTHSLPRPAAVVPRLMOP 360
DB 301 YNLTSDVSGCTLVCPRLHNOEVTAEEDGTORCEKSPCARGTHSLPRPAAVVPRLMOP 360
OY 361 GPAHPVLSFLRPSMDVSAFYSLPLAPLSPTSPVSPVSGRPPDDAHVAVNLSTYEG 419
DB 361 GPAHPVLSFLRPSMDVSAFYSLPLAPLSPTSPVSPVSGRPPDDAHVAVNLSTYEG 419

RESULT 12
AAE09211
ID AAE09211 standard: Protein: 419 AA.
XX
AC AAE09211;
XX
DT 15-NOV-2001 (first entry)
XX
DE Human p68HER-2 generic protein variant 9.
XX
KW HER-2; herstatin; antagonist; extracellular domain; ECD; Hereceptin;
XX
```

KM solid tumour; cancer; polymorphism; cytostatic; gene therapy;
 KW p68HER-2; ECDI11a; variant.
 XX
 OS Homo sapiens.
 PH Key Location/Qualifiers
 FT Region 1..340
 FT /note= "Identical to N-terminal region of p185HER-2"
 FT Domain 341..419
 FT /label= "ECDI11a variant"
 FT /note= "Extracellular domain I11a variant".
 FT Misc-difference 124
 FT /note= "Represented as Agn in the parent sequence shown
 FT in the specification"
 FT Misc-difference 125
 FT /note= "Represented as Agn in the parent sequence shown
 FT in the specification"
 FT Misc-difference 342
 FT /label= Unknown
 FT Misc-difference 345
 FT /label= Unknown
 FT Misc-difference 346
 FT /label= Unknown
 FT Misc-difference 356
 FT /label= Unknown
 FT Misc-difference 358
 FT /label= Unknown
 FT Misc-difference 361
 FT /label= Unknown
 FT Misc-difference 376
 FT /label= Unknown
 FT Misc-difference 394
 FT /label= Unknown
 FT Misc-difference 404
 FT /note= "p68HER-2 generic sequence (AAE09181) xaa
 FT substituted with leu"
 FT Misc-difference 413
 FT /label= Unknown
 XX
 XX WO200161356-A1.
 PD 23-AUG-2001.
 XX
 XX 16-FEB-2001; 2001WO-US05327.
 PF
 XX 16-FEB-2000; 2000US-0506079.
 PR
 XX (UYOR-) UNIV OREGON HEALTH SCI.
 PA
 XX Clinton G, Henner WD, Evans A;
 PI
 XX
 XX
 XX
 DR MPI; 2001-529934/58.
 XX
 XX
 PT New polypeptide, which binds to the extracellular domain of HER-2 for
 XX the treatment of hard tumors -
 XX
 XX
 PS
 XX
 XX
 CC The invention relates to novel HER-2 (herstatin-2) antagonist
 CC particularly a polypeptide that binds to the extracellular domain (ECD)
 CC of HER-2 at a site that is different from the binding site of humanised
 CC antibody, Herceptin, at an affinity of at least 10⁸. The present
 CC invention is based upon the initial discovery of an alternative HER-2
 CC mRNA transcript with 274 bp insert of intron 8. The translation product
 CC of the alternative transcript is a truncated HER-2 protein designated
 CC p68HER-2 which lacks the transmembrane and intracellular domains of
 CC p185HER-2 but contains ECD 1, II of the p185HER-2 and the novel ECDI11a.
 CC The ECDI11a-containing polypeptides bind tightly to, and thus antagonise
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
 CC nucleic acids encoding these are useful to treat, diagnose and identify
 CC solid tumours. The present sequence is human p68HER-2 generic protein
 CC containing ECDI11a variant sequence.
 CC Note: The present sequence is not shown in the specification but is

CC derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
 CC sequence listing (AAE09181).
 CC
 XX
 SQ Sequence 419 AA:
 Query Match 94.8%; Score 2168; DB 22; Length 419;
 Best Local Similarity 95.5%; Pred. No. 8.2e-168;
 Matches 400; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
 QY 1 MEALALCRWGLLALLPRGAASIQVCTGTDMLRLPASPETNIDMLRNLYOGCQVVGNL 60
 DB 1 MEALALCRWGLLALLPRGAASIQVCTGTDCKLRLPASPETNIDMLRNLYOGCQVVGNL 60
 QY 61 ELTYLPNASTLFODIQEYGVYLAHNOVROVPLQRIYVGTOLFEDNVALAVLDNG 120
 DB 61 ELTYLPNASTLFODIQEYGVYLAHNOVROVPLQRIYVGTOLFEDNVALAVLDNG 120
 QY 121 DPLNTPPVYGAASPGGLREQLRLTEILKGVLIQRNPOLCYQDTILMKDIFHKNNOLA 180
 DB 121 DPLXHTTPVYGAASPGGLREQLRLTEILKGVLIQRNPOLCYQDTILMKDIFHKNNOLA 180
 QY 181 LTLIDNRSRACHPCSPMCKSGKSGWESSDCLRTYVAGGCARCKPLPTDCCHEQC 240
 DB 181 LTLIDNRSRACHPCSPCKSGKSGWESSDCLRTYVAGGCARCKPLPTDCCHEQC 240
 QY 241 AAGCTGPKHSDDLACLFHNSGICELHCPALVYNTDTEFSMNPGRGYRFGASCVTACP 300
 DB 241 AAGCTGPKHSDDLACLFHNSGICELHCPALVYNTDTEFSMNPGRGYRFGASCVTACP 300
 QY 301 YNYLSTDVSGSCTLVCPRLHNOEVTAEQDQRCESKRPCARGTHSLPRPAVVPPLRMQP 360
 DB 301 YNYLSTDVSGSCTLVCPRLHNOEVTAEQDQRCESKRPCARGTHSLPRPAVVPPLRMQP 360
 QY 361 GPHAPVLSFLRPSMDLVSAFYSLPLAPLSPTSVIPISVYSGRPPDAHYAVNLSTREG 419
 DB 361 XPAHPVLSFLRPSMDVSAFYSLPLAPLSPTSVIPISVYSGRGLDPAHYAVXLSTREG 419
 RESULT 13
 AAE09204
 ID AAE09204 standard; Protein; 419 AA.
 XX
 XX AAE09204;
 AC
 XX
 DT 15-NOV-2001 (first entry)
 XX
 DE Human p68HER-2 generic protein variant 2.
 XX
 DE
 XX
 KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;
 KW p68HER-2; ECDI11a; variant.
 XX
 XX
 OS Homo sapiens.
 PH Key Location/Qualifiers
 FT Region 1..340
 FT /note= "Identical to N-terminal region of p185HER-2"
 FT Domain 341..419
 FT /label= "ECDI11a variant"
 FT /note= "Extracellular domain I11a variant"
 FT Misc-difference 124
 FT /note= "Represented as Agn in the parent sequence shown
 FT in the specification"
 FT Misc-difference 125
 FT /note= "Represented as Agn in the parent sequence shown
 FT in the specification"
 FT Misc-difference 342
 FT /label= Unknown
 FT Misc-difference 345
 FT /label= Unknown
 FT Misc-difference 346
 FT /note= "p68HER-2 generic sequence (AAE09181) xaa
 FT substituted with Pro"
 FT Misc-difference 346
 FT /label= Unknown

```

FT Misc-difference 356
FT /label= Unknown
FT Misc-difference 358
FT /label= Unknown
FT Misc-difference 361
FT /label= Unknown
FT Misc-difference 376
FT /label= Unknown
FT Misc-difference 394
FT /label= Unknown
FT Misc-difference 404
FT /label= Unknown
FT Misc-difference 413
FT /label= Unknown
FT WO200161356-A1.
FT 23-AUG-2001.
FT 16-FEB-2001; 2001WO-US05327.
FT 16-FEB-2000; 2000US-0506079.
FT (UYOR-) UNIV OREGON HEALTH SCI.
FT Clinton G, Henner WD, Evans A;
FT MPI; 2001-529934/58.
XX New polypeptide, which binds to the extracellular domain of HER-2 for
XX the treatment of hard tumors -
XX
XX Example 11: Page -: 61pp: English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Hereceptin, at an affinity of at least 10-8. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIia.
XX The ECDIIia-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 generic protein
XX containing ECDIIia variant sequence.
XX Note: The present sequence is not shown in the specification but is
XX derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
XX sequence listing (AAE09181).
XX
XX Sequence 419 AA:
XX
Query Match 94.8%; Score 2167; DB 22; Length 419;
Best Local Similarity 95.5%; Pred. No. 9,9e-168;
Matches 400; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```

```

QY 241 AAGCTGPKHSDCLACLFPHNHSIGICELCPALVTYNTDPESMPNPEGRTYFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFPHNHSIGICELCPALVTYNTDPESMPNPEGRTYFGASCVTACP 300
QY 301 YNVLSTDVSGCTLVCPJLHNOEVTAEEDGTGRCCKSPCARGTSHLLPRPAVPVPLRMQP 360
DB 301 YNVLSTDVSGCTLVCPJLHNOEVTAEEDGTGRCCKSPCARGTSHLLPRPAVPVPLRMQP 360
QY 361 GRAHPVLSFLRSPMDVSAFYSLPLAPLSPTSPVSPVSGRGPDPDAHVAVNLSRYEG 419
DB 361 XPAHPVLSFLRSPMDVSAFYSLPLAPLSPTSPVSPVSGRGPDPDAHVAVNLSRYEG 419

RESULT 14
AAE09216
ID AAE09216 standard; Protein; 419 AA.
XX
XX AAE09216;
XX
XX 15-NOV-2001 (first entry)
XX
XX Human p68HER-2 generic protein variant 15.
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Hereceptin;
XX solid tumour; cancer; polymorphism; cytostatic; gene therapy;
XX p68HER-2; ECDIIia; variant.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH 1..340
XX FT /note="Identical to N-terminal region of p185HER-2"
XX FT Domain 341..419
XX FT /label= ECDIIia_variant
XX FT /note="Extracellular domain Iia"
XX FT Misc-difference 124
XX FT /note="Represented as Agn in the sequence shown in
XX FT the specification"
XX FT Misc-difference 125
XX FT /note="Represented as Agn in the sequence shown in
XX FT the specification"
XX FT Misc-difference 342
XX FT /label= Unknown
XX FT Misc-difference 345
XX FT /label= Unknown
XX FT Misc-difference 346
XX FT /label= Unknown
XX FT Misc-difference 356
XX FT /label= Unknown
XX FT Misc-difference 357
XX FT /label= Unknown
XX FT Misc-difference 358
XX FT /label= Unknown
XX FT Misc-difference 361
XX FT /label= Unknown
XX FT Misc-difference 371
XX FT /label= Unknown
XX FT Misc-difference 376
XX FT /label= Unknown
XX FT Misc-difference 394
XX FT /label= Unknown
XX FT Misc-difference 404
XX FT /label= Unknown
XX FT Misc-difference 413
XX FT /label= Unknown
XX FT /note="p68HER-2 generic sequence (AAE09183) Xaa
XX substituted with Asn"
XX
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US05327.

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:31:51 ; Search time 24.3996 Seconds
(without alignments)
1650.860 Million cell updates/sec

Title: US-09-234-208B-2

Perfect score: 2287
Sequence: 1 MEALALCRWGLLLALLPPGA.....VGRGPDPAHVAVNLRYEG 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1878	82.1	1255	1 A24571	protein-tyrosine k
2	1608.5	70.3	1260	1 TVRTNU	protein-tyrosine k
3	1571	68.7	1254	2 I48161	p-185 precursor -
4	796.5	34.8	1223	1 TVCHLV	epidermal growth f
5	794	34.7	527	2 A42032	epidermal growth f
6	793	34.7	1210	1 GCHUE	epidermal growth f
7	789	34.5	1210	2 A53183	epidermal growth f
8	781.5	34.2	644	2 A35325	epidermal growth f
9	775	33.9	1308	2 A47253	epidermal growth f
10	735.5	32.2	1342	2 A36223	kinase-related tra
11	692	30.3	1339	2 JC4387	epidermal growth f
12	681.5	29.8	1166	1 S06142	protein-tyrosine k
13	575.5	25.2	843	2 A27131	epidermal growth f
14	419	18.3	1323	2 E88257	epidermal growth f
15	419	18.3	1374	2 S70712	protein-tyrosine k
16	416	18.2	1369	2 S70713	protein-tyrosine k
17	415	18.1	1330	1 G0FFE	epidermal growth f
18	363.5	15.9	366	2 D45558	epidermal growth f
19	363.5	15.9	1717	1 A45558	epidermal growth f
20	342.5	15.0	1363	2 T43220	insulin-like growt
21	331	14.5	333	2 B45558	epidermal growth f
22	331	14.5	342	2 C45558	epidermal growth f
23	294	12.9	2101	2 A57245	insulin receptor (
24	294	12.9	2148	1 A56081	insulin receptor -
25	291	12.7	1477	2 T18534	protein-tyrosine k
26	278	12.2	1300	2 A36502	insulin receptor-t
27	272	11.9	540	2 B47417	insulin receptor-t
28	269.5	11.8	1607	2 T43212	insulin-like growt
29	263.5	11.5	1382	1 INHUR	insulin receptor p

30	263.5	11.5	1383	2 A36080	insulin receptor p
31	261	11.4	1372	2 A34157	insulin receptor p
32	258	11.3	1390	2 T30346	insulin receptor -
33	257.5	11.3	1367	1 IGHUR1	insulin-like growt
34	252.5	11.0	1371	2 A33837	insulin-like growt
35	249	10.9	1268	2 B36502	insulin receptor-t
36	235	10.3	329	2 A48805	insulin-like growt
37	226.5	9.9	183	2 JH0803	tyrosine kinase re
38	203	8.9	1846	2 T42047	insulin receptor h
39	160.5	7.0	1299	2 T43251	furin (EC 3.4.21.1
40	144	6.3	1548	2 S34583	serine proteinase
41	142	6.2	1959	1 AGRT	agrin - rat
42	136.5	6.0	1111	2 T26972	hypothetical prote
43	133.5	5.8	915	1 A48225	subtilisin-like pr
44	132	5.8	1574	2 T13954	MEGF6 protein - ra
45	131.5	5.7	915	2 B46225	probable proteol

ALIGNMENTS

RESULT 1
A24571
protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human
N:Alternate names: c-erb-B-2 protein precursor; kinase-related transforming protein e
C:Species: Homo sapiens (man)
C:Date: 25-Oct-1987 #sequence revision 06-Dec-1996 #text change 11-Jun-1999
C:Accession: A24571; A25491; A44188; B44188; I59509; I57622
R:Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T
Nature 319, 230-234, 1986
A>Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growt
A:Reference number: A24571; MUID:8618663; PMID:3003577
A:Accession: A24571
A:Molecule type: mRNA
A:Residues: 1-1255 <YAM>
R:Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.
Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985
A>Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epid
A:Reference number: A25491; MUID:86016729; PMID:2995967
A:Accession: A25491
A:Molecule type: DNA
A:Residues: 737-1031 <SEM>
A:Cross-references: GB:M11767; NID:9182163; PIDN:AAA35808.1; PID:9553282
R:Cousens, L.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg
Science 230, 1132-1139, 1985
A>Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chro
A:Reference number: A44188; MUID:86070181; PMID:2995974
A:Accession: A44188
A:Molecule type: DNA
A:Residues: 740-910 <COU1>
A:Cross-references: GB:M12036; NID:9183988; PIDN:AAA35978.1; PID:9183989
A:Accession: B44188
A:Molecule type: mRNA
A:Residues: 1-517, 'RALL', 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>
A:Cross-references: GB:M11730; NID:9183986
R:King, C.R.; Kraus, M.H.; Aaronson, S.A.
Science 229, 974-976, 1985
A>Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.
A:Reference number: I59509; MUID:85272597; PMID:2992089
A:Accession: I59509
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 832-909 <REX>
A:Cross-references: NID:9459807; PIDN:AAA35809.1; PID:9459808
R:Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D.
Mol. Cell. Biol. 7, 2597-2601, 1987
A>Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptio
A:Reference number: I57622; MUID:87286898; PMID:3039351
A:Accession: I57622
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-191 <TAL>

A:Cross-references: GB:M16792; NID:G183983; PIDN:AAA58637.1; PID:9553332
 C:Comment: Amplification and overexpression of this erbB-related gene occurs in about 30
 C:Genetics:
 A:Gene: GDB:ERBB2; NGL: NEU; HER-2
 A:Cross-references: GDB:120613; OMIM:164870
 A:Map position: 17q21.1-19q21.1
 A:Introns: 25/1; 75/3; 147/1; 883/3
 A>Note: the list of introns is incomplete
 C:Function:
 A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospholase
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:12-1255/Product: protein-tyrosine kinase erbB2 #status predicted <MAT>
 F:22-653/Domain: extracellular #status predicted <EXT>
 F:70-304/Domain: EGF receptor extracellular domain repeat <EE1>
 F:395-605/Domain: EGF receptor extracellular domain repeat <EE2>
 F:654-675/Domain: transmembrane #status predicted <TM>
 F:676-1255/Domain: intracellular #status predicted <INT>
 F:718-983/Domain: protein kinase homology <KIN>
 F:726-734/Region: protein kinase ATP-binding motif
 F:68-124,187,259,530,571,629/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F:753/Active site: Lys #status predicted
 F:1133,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)
 Query Match 82.1%; Score 1878; DB 1; Length 1255;
 Best local Similarity 83.0%; Pred. No. 5,9e-121;
 Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 METALCRMGILLALLPRGAATVOCTGDMKRLRASPETHIDMLRHLXGCGVVGNL 60
 DB 1 METALCRMGILLALLPRGAATVOCTGDMKRLRASPETHIDMLRHLXGCGVVGNL 60
 QY 61 ELTYLPTNLSLFDIOEVOGYVLLAHNOVROVRLQRLRYVGTOLFEDNYALAVLDNG 120
 DB 61 ELTYLPTNLSLFDIOEVOGYVLLAHNOVROVRLQRLRYVGTOLFEDNYALAVLDNG 120
 QY 121 DPLNNTPTVTGASPGGLRELQSLTEILKGVLIQRNPOLCYODITLMKDIFHKNNOLA 180
 DB 121 DPLNNTPTVTGASPGGLRELQSLTEILKGVLIQRNPOLCYODITLMKDIFHKNNOLA 180
 QY 181 LTLIDTNRSPACHPCSPMKSGRCWGESSESDQSLTRTVACGAGCARCKGLPTDCHEQC 240
 DB 181 LTLIDTNRSPACHPCSPMKSGRCWGESSESDQSLTRTVACGAGCARCKGLPTDCHEQC 240
 QY 241 AAGCTPRKHSDDLACLFHNSGICELHCPALVTYNTDFESMNPBGRYTFGASCTTAC 300
 DB 241 AAGCTPRKHSDDLACLFHNSGICELHCPALVTYNTDFESMNPBGRYTFGASCTTAC 300
 QY 301 YNYLSTDVSGCTLVCPRLHNOEYTAEDGTORCEKSKPCAR----GTHSLPRPAAVPY 355
 DB 301 YNYLSTDVSGCTLVCPRLHNOEYTAEDGTORCEKSKPCARVCGVGLGMEHRLREVAATSAN 360
 QY 356 LRMQPG--PAHRYLSTLRPSMDLVSAFYSLPLAPLSPTSVPT-----SPVSGRGPD 405
 DB 356 IOEFACCKRIFGSLAPLPESFDGPASNT---APLQPEOLQVETLEITGYLYISAMPD 417
 QY 406 --PDAHVAVNLRYEG 419
 DB 418 SLPDLSVFQNLQVIRG 433

RESULT 2
 TYRTNU
 Protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 11-Jun-1999
 C:Accession: A24562; A61204
 R:Bargmann, C.I.; Hung, M.C.; Weinberg, R.A.
 Nature 319, 226-230, 1986
 A:Title: The neu oncogene encodes an epidermal growth factor receptor-related protein.

A:Reference number: A24562; MUID:86118662; PMID:3945311
 A:Accession: A24562
 A:Molecule type: mRNA
 A:Residues: 1-1260 <BAR>
 A:Cross-references: EMBL:X03362; NID:956745; PIDN:CAA27059.1; PID:956746
 R:Masui, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohe
 Carcinogenesis 12, 1975-1978, 1991
 A:Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals n
 2-thiazolyl]formamide or N-methyl-N-nitrosourea.
 A:Reference number: A61204; MUID:92035293; PMID:1682063
 A:Accession: A61204
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 637-663, 'V', 665-702 <MAS>
 A>Note: authors translated the codon GCA for residue 25 as Val
 C:Genetics:
 A:Gene: neu
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phos
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-1260/Product: protein-tyrosine kinase neu #status predicted <MAT>
 F:658-680/Domain: transmembrane #status predicted <TM>
 F:723-988/Domain: protein kinase homology <KIN>
 F:731-739/Region: protein kinase ATP-binding motif
 F:1,191,263,535,576,634/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:691/Binding site: phosphate (Thr) (covalent) #status predicted
 F:758/Active site: Lys #status predicted
 F:882,1227,1253/Binding site: phosphate (Tyr) (covalent) #status predicted
 Query Match 70.3%; Score 1608.5; DB 1; Length 1260;
 Best local Similarity 85.0%; Pred. No. 1.8e-102;
 Matches 294; Conservative 15; Mismatches 36; Indels 1; Gaps 1;

QY 1 METALCRMGILLALLPRGAATVOCTGDMKRLRASPETHIDMLRHLXGCGVVGNL 60
 DB 4 METALCRMGILLALLPRGAGTVOCTGDMKRLRASPETHIDMLRHLXGCGVVGNL 63
 QY 61 ELTYLPTNLSLFDIOEVOGYVLLAHNOVROVRLQRLRYVGTOLFEDNYALAVLDNG 120
 DB 64 ELTYLVANMSLSLFDIOEVOGYMLLAHNOVKRVRQLRLRYVGTOLFEDNYALAVLDNR 123
 QY 121 DPLNNTPTVT-GASPGGLRELQSLTEILKGVLIQRNPOLCYODITLMKDIFHKNNOL 179
 DB 124 DPQDNVAASPTGRTPRGLRELQSLTEILKGVLLIRGNPOLCYODIMKDVFRKNNOL 183
 QY 181 LTLIDTNRSPACHPCSPMKSGRCWGESSESDQSLTRTVACGAGCARCKGLPTDCHEQC 239
 DB 184 APVDIDTNRSPACHPCSPMKSGRCWGESSESDQSLITGITISGCAKCKGRPLTDCHEQ 243
 QY 240 CAAGCTPRKHSDDLACLFHNSGICELHCPALVTYNTDFESMNPBGRYTFGASCTTAC 299
 DB 244 CAAGCTPRKHSDDLACLFHNSGICELHCPALVTYNTDFESMNPBGRYTFGASCTTAC 303
 QY 300 PNYLSTDVSGCTLVCPRLHNOEYTAEDGTORCEKSKPCARAGNSL 345
 DB 304 PNYLSTDVSGCTLVCPRLHNOEYTAEDGTORCEKSKPCARVCYGL 349

RESULT 3
 I48161
 P-185 precursor - golden hamster
 C:Species: Mesocricetus auratus (golden hamster)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999
 C:Accession: I48161
 R:Nakamura, T.; Ushijima, T.; Ishizaka, Y.; Nagao, M.; Arai, M.; Yamazaki, Y.; Ishika
 Gene 140, 251-255, 1994
 A:Title: Cloning and activation of the Syrian hamster neu proto-oncogene.
 A:Reference number: I48161; MUID:94193007; PMID:7908275
 A:Accession: I48161
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1254 <RES>
 A:Cross-references: GB:D16295; NID:9493236; PIDN:BA03801.1; PID:9747595

[illegible]

A:Residues: 713-964 <LIN>
A:Experimental source: epidermoid carcinoma cell line A431
R.Stimmen, P.A.; Gope, M.L.; Schulz, T.Z.; Wright, D.A.; Carpenter, G.; O'Walley, B.W.
Biochim. Biophys. Res. Commun. 124, 125-132, 1984
A:Reference number: A23062; PMID:85046483; PMID:6093780
A:Accession: A23062
A:Molecule type: mRNA
A:Residues: 1028-1210 <STM>
R.Weber, W.; Gill, G.N.; Speiss, J.
Science 224, 294-297, 1984
A:Reference number: A05281; PMID:84172183; PMID:6324343
A:Accession: A05281
A:Molecule type: protein
A:Residues: 25-30.'S'/32-51.454-467 <WEB>
R.Russo, M.M.; Lukas, T.J.; Cohen S.; Staresos, J.V.
J. Biol. Chem. 260, 5205-5208, 1985
A>Title: Identification of residues in the nucleotide binding site of the epidermal growth factor receptor
A:Reference number: A60143; PMID:85182650; PMID:2985580
A:Accession: A60143
A:Molecule type: protein
A:Residues: 740-744,'X'/746-747 <RUS>
R.Mrozowski, B.; Mosig, G.; Cohen, S.
Nature 309, 270-273, 1984
A>Title: ATP-stimulated interaction between epidermal growth factor receptor and superoxide dismutase
A:Reference number: A38023; PMID:8419154; PMID:6325948
A:Contents: annotation; receptor activity
A>Note: The EGF receptor (and other tyrosine kinases) can nick double-stranded DNA
R.Chen, W.S.; Lazar, C.S.; Lund, K.A.; Weisn, J.B.; Chang, C.P.; Walton, G.M.; Der, C.G.
Cell 59, 33-43, 1989
A>Title: Functional independence of the epidermal growth factor receptor from a domain involved in ligand-induced internalization
A:Reference number: A3331; PMID:9003233; PMID:2790960
A:Contents: annotation; internalization signal
C.Comment: Binding of EGF to the receptor leads to internalization of the EGF-receptor complex
C.Genetics:
A:Gene: GDB:EGFR
A:Cross-references: GDB:120610; OMIM:131550
A:Map position: 7p12.3-7p12.1
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C.Keywords: Arp; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphorylation; signal transduction; status predicted
F:1-24/Domin: signal sequence #status predicted <SIG>
F:25-1210/Product: EGF receptor #status predicted <MAT>
F:25-645/Domai: extracellular #status predicted <EXT>
F:75-300/Domai: EGF receptor extracellular domain repeat <EB1>
F:390-600/Domai: EGF receptor extracellular domain repeat <BE2>
F:646-668/Domai: transmembrane #status predicted <TMW>
F:669-1210/Domai: intracellular #status predicted <INT>
F:710-975/Domai: protein kinase homology <KIN>
F:718-726/Region: protein kinase Arp-binding motif
F:999-1046/Region: coated-pit mediated internalization signal
F:1047-1210/Region: inhibitory
F:128.175.352.413.444.528.603/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:745/Active site: lys #status experimental

Query Match 34.7%, Score 793; DB 1; Length 1210;
Best Local Similarity 45.3%; Pred. No. 1.3e+46;
Matches 151; Conservative .50; Mismatches 118; Indels 14; Gaps 4;

Qy	308	VGSGTLYCPHLNDEVTADGTQRCCEKSKSCPKCAR	340
Db	304	HGSCVCRACGADSYEM-EDDGVRKCKCKEGEGCRK	335
Qy	248	KHSQSLACIAFHNHSGIIECLHPALVTNTPTFSPMPDEGRARYEGASCVYACAPNYLSTD	307
Db	244	RESQIVCRKRDRDCAITCKDPCPLMLNLNPTTYQMDVNPBESKYSGFATCYVKKCPRNTVYTD	303

RESULT 7

epidermal growth factor receptor precursor mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence,revision 06-Jan-1995 #text_change 18-Jun-1999
C:Accession: A53183; A43818; S24942; A28941; S45325; I49643
R:Luettke, N.C.; Phillips, H.K.; Qiu, T.H.; Copeland, N.G.; Earp, H.S.; Jenkins, N.A.;
Genes Dev. 8, 399-413, 1994
A>Title: The mouse waved-2 phenotype results from a point mutation in the EGF receptor
A:Reference number: A53183; MUID:94170986; PMID:812525
A:Accession: A53183
A:Molecule type: mRNA
A:Residues: 1-1210 <LUE>
A:Cross-references: GB:003425
R:Avivi, A.; Lax, I.; Ullrich, A.; Schlessinger, J.; Givol, D.; Morse, B.
Oncogene 6, 673-676, 1991
A>Title: Comparison of EGF receptor sequences as a guide to study the ligand binding site
A:Reference number: A43818; MUID:91232866; PMID:2030916
A:Accession: A43818
A:Molecule type: mRNA
A:Residues: 1-714 <AVI>
A:Cross-references: GB:X59698
R:Eisinger, D.P.; Serrero, G.
submitted to the EMBL Data Library, June 1992
A:Reference number: S24942
A:Accession: S24942
A:Molecule type: mRNA
A:Residues: 969-971, 'K', 973-1115, 'D' <EIS>
A:Cross-references: EMBL:Z12608
R:Heisermann, G.J.; Gill, G.N.
J. Biol. Chem. 263, 13152-13158, 1988
A>Title: Epidermal growth factor receptor threonine and serine residues phosphorylated
A:Reference number: A28941; MUID:88330814; PMID:3138233
A:Accession: A28941
A:Molecule type: protein
R:Hibbs, M.L.; Dunn, A.R.; Alexander, W.S.
submitted to the EMBL Data Library, April 1994
A:Description: The complete cDNA sequence of the Mouse Epidermal Growth Factor Receptor
A:Reference number: S45325
A:Accession: S45325
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-971, 'K', 973-1210 <VER>
A:Cross-references: EMBL:X78987; NID:9488830; PIDN:CAA55587.1; PID:9488831
R:Paria, B.C.; Das, S.K.; Andrews, G.K.; Dey, S.K.
Proc. Natl. Acad. Sci. U.S.A. 90, 55-59, 1993
A>Title: Expression of the epidermal growth factor receptor gene is regulated in mouse
A:Reference number: I49643; MUID:93126380; PMID:7678348
A:Accession: I49643
A>Status: translated from GB/EMBL/DDBT
A:Molecule type: mRNA
A:Residues: 12-20,22-132 <RES>
A:Cross-references: GB:U06864; NID:g193001; PIDN:AA55029.1; PID:g567201
C:Genetics:
A:Gene: EGFR
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: ATP; growth factor receptor; kinase-related transforming protein; phosphoprotein
F:1-24/Domain: signal sequence #status predicted <SIG>
F:648-670/Domain: transmembrane #status predicted <TM>
F:712-977/Domain: protein kinase homology <KIN>
F:720-728/Region: protein kinase ATP-binding motif
F:680,695/Binding site: phosphate (Thr) (covalent) #status experimental
F:697,1070,1071/Binding site: phosphate (Ser) (covalent) #status experimental

Db 236 CAAGCTGPRESDCLVCHRRPREATCKDTCPPMLVNPPTTYQMDVNEGKVSFGATCVKRC 295
 QY 300 PNYLSTDVSGCTLVCPRLHNOEVTAEQDROCKSKPCAR 340
 Db 296 PKNYVTDHSGCVACGPPDYEV-EEQVSKCKCKDCGPKR 335

RESULT 9
 A47253
 epidermal growth factor receptor, HER4 - human

C:Species: Homo sapiens (man)
 C:Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999
 C:Accession: A47253
 R:Plowman, G.D.; Clouscou, J.M.; Whitney, G.S.; Green, J.M.; Carlton, G.W.; Foy, L.; NE
 Proc. Natl. Acad. Sci. U.S.A. 90, 1746-1750, 1993
 A:Title: Ligand-specific activation of HERK/p180erbB4, a fourth member of the epidermal
 A:Reference number: A47253; MUID:93189574; PMID:8383326
 A:Accession: A47253
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: nucleic acid
 A:Residues: 1-1308 <PIG>
 A:Cross-references: GB:107868; NID:9337359; PIDN:AAB59446.1; PID:9337360
 A:Note: sequence extracted from NCBI backbone (NCBIP:126842)
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; growth factor receptor
 F:716-981/Domain: protein kinase homology <KIN>
 F:724-732/Region: protein kinase ATP-binding motif

Query Match 33.9%; Score 775; DB 2; Length 1308;
 Best Local Similarity 45.1%; Pred. No. 2.4e-45;
 Matches 157; Conservative 45; Mismatches 124; Indels 22; Gaps 7;

QY 9 WGLLALLPPGAA---STQVCTGTDMLRLPASPETHLMDLRHLVGGGVQVGNLELY 64
 Db 8 WWSVLLVAGVYQPSDSQSCAGTENKLSLSLEQVRLRKTYENCEVWGNLEITS 67
 QY 65 LPTNASLFLQDIOEVQVYLIAHNOVROVPLORLRIYRGTOLEFEDNALAVLDNGDPLN 124
 Db 68 IEHNRDLFSRVYEVGLVALNOFRYLPENLRIRGTLYEDRYALAIPLNRYKDG 127
 QY 125 NTPPYVGASPGRLRELQSLSTFPIELKGVLIQNPOLCYQDTILMKDIFHKNNQALTL 184
 Db 128 NF-----GLQELGLNLNTEILNGVYVDQNKFLCYADTILMODIVRNWPSNLTLLV 178
 QY 185 DYNRSRACHPCSPMGKSGRSGESSEDCQSLRTVCAGGC-ARCKGRLPTDCHCEGCAAG 243
 Db 179 STNGSSGCGRCHKSGTG-KCMGPTENHCOTLRTVCACBQCDGCGYPRYSDCHRECAAG 237
 QY 244 CTGPKHSDCLAHFNHSGICELHCPALVYNTDFESMPNPEGRYTFGASCVTACPYNY 303
 Db 238 CSGPKPTDFACMNFNDSCAVYQCQPTFVYNPPTFQLEHNFNAKYTGAFCKPCPNHF 297
 QY 304 LSTDVQSCVLVCPRLHNOEVTAEQDROCKSKPCAR-----GTHSL 346
 Db 298 V-VDSSSCVACRPPSKMEV-EENGIMKRCPCDIDICPKACDGIQGLSM 343

RESULT 10
 A36223
 kinase-related transforming protein (erbB3) (BC 2.7.1.-) precursor - human
 C:Species: Homo sapiens (man)
 C:Date: 04-Oct-1991 #sequence_revision 13-Jan-1993 #text_change 17-Nov-2000
 C:Accession: A36223; I59164
 R:Kraus, M.H.; Issing, W.; Mikki, T.; Popescu, N.C.; Aaronson, S.A.
 Proc. Natl. Acad. Sci. U.S.A. 86, 9193-9197, 1989
 A:Title: Isolation and characterization of ERB3, a third member of the ERBB/epidermal g
 A:Reference number: A36223; MUID:90083234; PMID:2687875
 A:Accession: A36223
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1342 <KRA>
 A:Cross-references: GB:M29366
 R:Plowman, G.D.; Whitney, G.S.; Neubauer, M.G.; Green, J.M.; McDonald, V.L.; Todaro, G.J

Proc. Natl. Acad. Sci. U.S.A. 87, 4905-4909, 1990
 A:Title: Molecular cloning and expression of another epidermal growth factor receptor
 A:Reference number: I59164; MUID:90311312; PMID:2164210
 A:Accession: I59164

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-559, 'G', '561-957', 'F', '959-1063', 'G', '1065-1342 <RES>
 A:Cross-references: GB:M34309; NID:9183990; PIDN:AAA35979.1; PID:9306841
 C:Genetics:
 A:Gene: GDB:ERBB3; HER3
 A:Cross-references: GDB:119880; OMIM:190151
 A:Map position: 12q13-12q13
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
 C:Keywords: ATP; phosphotransferase
 F:707-972/Domain: protein kinase homology <KIN>
 F:715-723/Region: protein kinase ATP-binding motif

Query Match 32.2%; Score 735.5; DB 2; Length 1342;
 Best Local Similarity 44.0%; Pred. No. 1.3e-42;
 Matches 150; Conservative 49; Mismatches 121; Indels 21; Gaps 9;

QY 10 GLLALLPPGAA--STQVCTGTDMLRLPASPETHLMDLRHLVGGGVQVGNLELYPT 67
 Db 11 GLFSLARSGEVNGSQAVCPGTLNGLSVGDAENQYQTLKYERCEVWGNLEITLGH 70
 QY 68 NLSLFLQDIOEVQVYLIAHNOVROVPLORLRIYRGTOLEFEDNALAVLDNGDPLNNT 127
 Db 71 NDLSTLQWIREVTGYVLVAMNEFTPLPRLRYRGTOYVYGRKALFPM-----LNNT 125
 QY 128 PVTGASPGRLRELQSLSTFPIELKGVLIQNPOLCYQDTILMKDIFHKNNQALTLIDTN 187
 Db 126 ---NSSHALROLRLTLQLEILSGVYIEKNDKLCIMDTIDRDIYRDD--AEIYVD 178
 QY 188 RSRACHPCSPMGKSGRSGESSEDCQSLRTVCAGGC-ARCKGRLPTDCHCEGCAAGCG 246
 Db 179 NGRSCPCHEVCKG-KCMGPGSEDCQTLTKTICAPQCNHCGPMPNQCHCEAGCGSG 237
 QY 247 PKHSDCLAHFNHSGICELHCPALVYNTDFESMPNPEGRYTFGASCVTACPYNYLST 306
 Db 238 PDQDGFACRHNHNDSCAVYQCQPTFVYNPPTFQLEHNFNAKYTGAFCKPCPNHFV 296
 QY 307 DVGSCVLVCPRLHNOEVTAEQDROCKSKPCAR---GTHS 344
 Db 297 DQTSVCVRACPPDKMEVD--KNGLMCEPCGCLPKACEGTGS 336

RESULT 11
 Jc4387
 epidermal growth factor receptor homolog precursor - rat
 N:Alternate names: ErbB3 protein; HER3 protein
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 17-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 13-Nov-1998
 C:Accession: Jc4387
 R:Heillyer, N.J.; Kim, H.H.; Greaves, C.H.; Sierke, S.L.; Koland, J.G.
 Gene 165, 279-284, 1995
 A:Title: Cloning of the rat ErbB3 cDNA and characterization of the recombinant protei
 A:Reference number: Jc4387; MUID:96096535; PMID:8522190
 A:Accession: Jc4387
 A:Molecule type: mRNA
 A:Residues: 1-1339 <HEL>
 A:Cross-references: GB:U29339; NID:9915389; PID:9915390
 A:Experimental source: liver
 A:Note: The authors translated the codon AAG for residue 369 as Thr and GTT for resid
 C:Comment: This protein is a functional heregulin receptor that transduces signals to
 C:Genetics:
 A:Gene: ErbB3
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
 C:Keywords: ATP; growth factor receptor; liver; phosphoprotein; transmembrane protein
 F:20-1339/Product: epidermal growth factor homolog
 F:640-659/Domain: transmembrane #status predicted <TM>
 F:705-970/Domain: protein kinase homology <KIN>
 F:713-721/Region: protein kinase ATP-binding motif

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:25:15 ; Search time 15.1446 seconds
(without alignments)
1147.512 Million cell updates/sec

Title: US-09-234-208b-2

Perfect score: 2287
Sequence: 1 MEIALCRWGLLALLPAGA.....VGRGPDPAHVAVNLSRYEG 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	187.8	82.1	1255	1 ERB2_HUMAN	P04626 homo sapien
2	1608.5	70.3	1257	1 ERB2_RAT	P06494 rattus norv
3	1571	68.7	1254	1 ERB2_MESAU	O60553 mesocricetu
4	796.5	34.8	703	1 EGFR_CHICK	P13387 gallus gall
5	793	34.7	1210	1 EGFR_HUMAN	P00533 homo sapien
6	789	34.5	1210	1 EGFR_MOUSE	O01279 mus musculu
7	779.5	34.1	1308	1 ERB4_RAT	O62956 rattus norv
8	775	33.9	1308	1 ERB4_HUMAN	O15303 homo sapien
9	735.5	32.2	1342	1 ERB3_HUMAN	P21860 homo sapien
10	698	30.5	1339	1 ERB3_RAT	O62799 rattus norv
11	681.5	29.8	1167	1 XMRK_XIPMA	P13388 xiphophorus
12	574.5	25.1	1426	1 EGFR_DROME	P04412 drosophila
13	419	18.3	1323	1 LT23_CABEL	P24348 caenorhabdi
14	342.5	15.0	1363	1 ILPR_BRALA	O02466 brachyost
15	294	12.9	2146	1 INSR_DROME	P09208 drosophila
16	291	12.7	1477	1 HTK7_HYDAT	O25197 hydra atten
17	290	12.7	1300	1 IRR_MOUSE	O9wt14 mus musculu
18	278	12.2	1300	1 IRR_CAVPO	P14617 cavia porce
19	270	11.8	581	1 IRR_RAT	O64716 rattus norv
20	269.5	11.8	1607	1 MIPR_LYMTST	O25410 lymphaea sta
21	269	11.8	1297	1 IRR_HUMAN	P14616 homo sapien
22	264.5	11.6	1382	1 INSR_HUMAN	P06213 homo sapien
23	263.5	11.5	1383	1 INSR_RAT	P15127 rattus norv
24	261	11.4	1372	1 INSR_MOUSE	O93105 mus musculu
25	258	11.3	1390	1 INSR_AEDAE	O93105 aedes aegypt
26	257.5	11.3	1367	1 IGR_HUMAN	P08069 homo sapien
27	252.5	11.0	1370	1 IGR_MOUSE	P24062 rattus norv
28	250.5	11.0	1373	1 IGR_MOUSE	O60791 mus musculu
29	169	7.4	1696	1 PKC5_BRACL	O94115 branchiost
30	144	6.3	1877	1 PKC5_MOUSE	O04552 mus musculu
31	142	6.2	1959	1 AGRI_RAT	P25304 rattus norv
32	131.5	5.7	913	1 PKC5_HUMAN	O92824 homo sapien
33	129.5	5.7	1680	1 FUR2_DROME	P30432 drosophila

34	128	5.6	830	1 SREC_HUMAN	O14162 homo sapien
35	124.5	5.4	1877	1 PKC5_RAT	P41413 rattus norv
36	123.5	5.4	937	1 PAC4_RAT	O63415 rattus norv
37	123	5.4	417	1 TR12_HUMAN	O93038 h tumor nec
38	120	5.2	498	1 TNFR_MOUSE	O60846 mus musculu
39	120	5.2	3695	1 LMAS_HUMAN	O15230 homo sapien
40	119.5	5.2	667	1 TS11_GIALA	O03185 glandia lam
41	119	5.2	2282	1 ZAN_RABIT	P57999 oryctolagus
42	118.5	5.2	551	1 LEM2_RABIT	P27113 oryctolagus
43	118.5	5.2	3718	1 LMAS_MOUSE	O61001 mus musculu
44	117	5.1	469	1 PROP_HUMAN	P27918 homo sapien
45	117	5.1	3075	1 LMAL_HUMAN	P25391 homo sapien

ALIGNMENTS

RESULT 1
ERB2_HUMAN
ID ERB2_HUMAN STANDARD: PRT; 1255 AA.
AC P04626;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-2 precursor (BC 2.7.1.112)
DE (P185erbB2) (NEU proto-oncogene) (C-erbB-2) (Tyrosine kinase-type cell
DE surface receptor HER2) (M19.19).
GN ERB2 OR HER2 OR NGL OR NEU.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=6118663; PubMed=3003577;
RA Yamamoto T., Ikawa S., Akiyama T., Semba K., Nomura N., Miyajima N.,
RA Saito T., Toyoshima K.;
RT "Similarity of protein encoded by the human c-erb-B-2 gene to
RT epidermal growth factor receptor.";
RL Nature 319:230-234(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=66070181; PubMed=2999974;
RA Cussens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A.,
RA McGrath J., Seeburg P.H., Libermann T.A., Schlessinger J.,
RA Francke U., Levinson A., Ullrich A.;
RT "Tyrosine kinase receptor with extensive homology to EGF receptor
RT shares chromosomal location with neu oncogene.";
RN Science 230:1132-1139(1985).
RN [3]
RP SEQUENCE OF 737-1031 FROM N.A.
RX MEDLINE=86016729; PubMed=2995967;
RA Semba K., Kamata N., Toyoshima K., Yamamoto T.;
RT "A v-erbB-related protooncogene, c-erbB-2, is distinct from the
RT c-erbB-1/epidermal growth factor-receptor gene and is amplified in a
RT human salivary gland adenocarcinoma.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).
RN [4]
RP VARIANTS VAL-654 AND VAL-655.
RX MEDLINE=93194196; PubMed=8095488;
RA Ehsani A., Low J., Wallace R.B., Wu A.M.;
RT "Characterization of a new allele of the human ERB2 gene by allele-
RT specific competition hybridization.";
RL Genomics 15:426-429(1993).
RN [5]
RP FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,
RP ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A
RP POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-
RP ALPHA AND AMPHIREGULIN.
CC CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS
CC (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
CC RESIDUES (BY SIMILARITY).
CC -1- POLYMORPHISM: THERE ARE FOUR ALLELES DUE TO THE VARIATIONS IN
CC POSITIONS 654 AND 655. ALLELE B1 (654-ILE-ILE-655) HAS A FREQUENCY
CC OF 0.782; ALLELE B2 (654-ILE-VAL-655) HAS A FREQUENCY OF 0.206;
CC ALLELE B3 (654-VAL-VAL-655) HAS A FREQUENCY OF 0.012.
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M11767; AAA35808.1; JOINED.
DR EMBL: M11761; AAA35808.1; JOINED.
DR EMBL: M11762; AAA35808.1; JOINED.
DR EMBL: M11763; AAA35808.1; JOINED.
DR EMBL: M11764; AAA35808.1; JOINED.
DR EMBL: M11765; AAA35808.1; JOINED.
DR EMBL: M11766; AAA35808.1; JOINED.
DR EMBL: M11730; AAA35808.1; JOINED.
DR EMBL: M12036; AAA35808.1; JOINED.
DR EMBL: M12036; AAA35808.1; JOINED.
DR PIR: A25491; A25491.
DR PIR: A25491; A25491.
DR HSSP: P11362; 1FGK.
DR Genew: HGNC:3430; ERBB2.
DR MIM: 164870; -.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR000719; Euk_kinase.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR001245; Tyr_kinase.
DR InterPro: IPR004019; YLP_motif.
DR Pfam: PF00069; kinase.1.
DR Pfam: PF00757; Furin-like.1.
DR Pfam: PF01030; Recep_L_domain.2.
DR Pfam: PF02757; YLP.2.
DR Prodom: PD000001; Euk_kinase.1.
DR SMART: SM00261; FU.3.
DR SMART: SM00219; TYRC.1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR.1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
KW transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW Polymorphism.
KW -----
FT CHAIN 1 21 POTENTIAL.
FT CHAIN 22 1255 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.
FT DOMAIN 22 652 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 653 675 POTENTIAL.
FT DOMAIN 676 1255 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 720 987 PROTEIN KINASE.
FT NP_BIND 726 734 ATP (BY SIMILARITY).
FT BINDING 753 753 ATP (BY SIMILARITY).
FT ACT_SITE 845 845 BY SIMILARITY.
FT DISULFID 195 204 BY SIMILARITY.
FT DISULFID 199 212 BY SIMILARITY.
FT DISULFID 220 227 BY SIMILARITY.
FT DISULFID 224 235 BY SIMILARITY.
FT DISULFID 236 244 BY SIMILARITY.
FT DISULFID 240 252 BY SIMILARITY.
FT DISULFID 255 264 BY SIMILARITY.
FT DISULFID 268 295 BY SIMILARITY.
FT DISULFID 299 311 BY SIMILARITY.
FT DISULFID 315 331 BY SIMILARITY.
FT DISULFID 334 338 BY SIMILARITY.
FT DISULFID 511 520 BY SIMILARITY.
FT DISULFID 515 528 BY SIMILARITY.
FT DISULFID 531 540 BY SIMILARITY.
FT DISULFID 544 560 BY SIMILARITY.

FT DISULFID 563 576 BY SIMILARITY.
FT DISULFID 567 584 BY SIMILARITY.
FT DISULFID 587 596 BY SIMILARITY.
FT DISULFID 600 623 BY SIMILARITY.
FT DISULFID 626 634 BY SIMILARITY.
FT DISULFID 630 642 BY SIMILARITY.
FT MOD_RES 1139 1139 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1248 1248 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 68 68 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 124 124 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 187 187 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 530 530 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 571 571 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 629 629 N-LINKED (GLCNAC...) (POTENTIAL).
FT VARIANT 654 654 I -> V.
FT VARIANT 654 654 /FTID-VAR_004077.
FT VARIANT 655 655 I -> V.
FT CONFLICT 1170 1170 /FTID-VAR_004078.
FT CONFLICT 1170 1170 P -> A (IN REF. 2).
SQ SEQUENCE 1255 AA: 137909 MW: 39899DDA04DC962 CRC64;
Query Match 82.1%; Score 1878; DB 1; Length 1255;
Best Local Similarity 83.0%; Pred. No. 6,7e-138;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
QY 1 MELALCRWGLLALPPGASSTQVCTGDMKLRPSPETHIDMLRHLVGGCGVGNL 60
Db 1 MELALCRWGLLALPPGASSTQVCTGDMKLRPSPETHIDMLRHLVGGCGVGNL 60
QY 61 ELTYLPTNASLSFLQDIQVGVYLIHNOVROVPLQRLIRVGTQLEEDNYALAVDNG 120
Db 61 ELTYLPTNASLSFLQDIQVGVYLIHNOVROVPLQRLIRVGTQLEEDNYALAVDNG 120
QY 121 DPLNNTPTVGVSGEGRLREILKGVLIQNPOLCTQDTITLMDIFHKNQOLA 180
Db 121 DPLNNTPTVGVSGEGRLREILKGVLIQNPOLCTQDTITLMDIFHKNQOLA 180
QY 181 LTLIDTRSRACHPCSPMGSRGCGSSSEDCSLPTVAGGARGKGLPDDCCEQC 240
Db 181 LTLIDTRSRACHPCSPMGSRGCGSSSEDCSLPTVAGGARGKGLPDDCCEQC 240
QY 241 AAGCTGPKHSDCLALFHNHSGICELHCPALVYNTDTPESMPNPEGRTYFGACVYACP 300
Db 241 AAGCTGPKHSDCLALFHNHSGICELHCPALVYNTDTPESMPNPEGRTYFGACVYACP 300
QY 301 YNYISTDVGSCITVCPPLHNOEYTAEDGTQCEKCKSKPCAR----GTHSLPPRAAVPV 355
Db 301 YNYISTDVGSCITVCPPLHNOEYTAEDGTQCEKCKSKPCARCYGLGMEHLREVARAVTSAN 360
QY 356 LRMQPG--PAHPVLSFLRPSMDLVSAFYSLLAPLSTSVPI-----SPVSYGSGPD 405
Db 361 IQEPAGCKKTFGSLALPESFDGPPASNT---APLOPEQLOVEFTLEITGYLYISAMPD 417
QY 406 --PDAAVAVNLRYEG 419
Db 418 SLDPDLSVFQNLQYIRG 433
RESULT 2
ERB2_RAT
ID ERB2_RAT STANDARD; PRT; 1257 AA.
AC P06484;
DT 01-JAN-1988 (Rel. 06, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)
DE (p185erbB2) (neu proto-oncogene) (C-erbB-2) (Epidermal growth factor
DE receptor-related protein).
GN ERBB2 OR NEU.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

RX	NCB1_taxid=10116;
RP	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Neuroblastoma;
RX	MEDLINE=86118662; PubMed=3945311;
RA	Bargmann C.I., Hung M.-C., Weinberg R.A.;
RT	"The neu oncogene encodes an epidermal growth factor receptor-related protein".
RL	Nature 319:226-230(1986).
RN	[2]
RN	SEQUENCE OF 852-905 FROM N.A.
RP	TISSUE=Sciatic nerve;
RX	MEDLINE=91222560; PubMed=2025425;
RA	Iai C., Lemke G.;
RT	"An extended family of protein-tyrosine kinase genes differentially expressed in the vertebrate nervous system.";
RL	Neuron 6:691-704(1991).
RN	[3]
RN	STRUCTURE BY NMR OF 650-668.
RX	MEDLINE=92155181; PubMed=1346763;
RA	Gullikc W.J., Bottomley A.C., Lotfs F.J., Doak D.G., Mulvey D., Newman R., Crumpton M.J., Stenberg M.J.E., Campbell I.D.;
RT	"Three dimensional structure of the transmembrane region of the proto-oncogenic and oncogenic forms of the new protein.";
RL	EMBO J. 11:43-48(1992).
CC	- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,
CC	ALTHOUGH NEEUREGLINS DO NOT INTERACT WITH IT ALONE. GP30 IS A POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-
CC	ALPHA AND AMPHIREGULIN.
CC	- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC	tyrosine phosphate.
CC	- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS.
CC	THE CONSTITUTIVELY ACTIVATED ONCOGENIC VARIANT FORMS A HOMODIMER.
CC	- SUBCELLULAR LOCATION: Type I membrane protein
CC	- PMR: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES (BY SIMILARITY).
CC	- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/isb-ch).
CC	or send an email to license@isb-sib.ch).
CC	-----
CC	EMBL: X03362; CAA27059.1; ALT_INIT.
DR	PIR: A24562; TVRRNU.
DR	HSP, P11362; IRGK.
DR	InterPro: IPR000494; EGFR_L_domain.
DR	InterPro: IPR000719; Euk_pkinase.
DR	InterPro: IPR002174; Furin-like.
DR	InterPro: IPR001245; Tyr_kinase.
DR	InterPro: IPR004019; YLP_motif.
DR	Pfam: PF00069; pkinase: 1.
DR	Pfam: PF00757; Furin-like: 1.
DR	Pfam: PF01030; Recep_L_domain: 2.
DR	Pfam: PF02757; YLP: 2.
DR	ProdDom: PD000001; Euk_pkinase: 1.
DR	SMART: SM00261; FU: 3.
DR	SMART: SM00219; TyrcK: 1.
DR	PROSITE: PS00107; PROTEIN_KINASE_ATP: 1.
DR	PROSITE: PS00109; PROTEIN_KINASE_TYR: 1.
DR	PROSITE: PS50011; PROTEIN_KINASE_DOM: 1.
KW	Transmembran; Glycoprotein; Multigene family; Receptor; Signal;
KW	Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW	Proto-oncogene; Disease mutation.
FT	SIGNAL 1 21
FT	CHAIN 1 21
FT	CHAIN 22 1257
FT	DOMAIN 22 654
FT	TRANSMEM 655 677
FT	DOMAIN 678 1257
FT	DOMAIN 159 369
FT	CYS-RICH
FT	POTENTIAL
FT	RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.
FT	EXTRACELLULAR (POTENTIAL).
FT	POTENTIAL.
FT	CYTOPLASMIC (POTENTIAL).
FT	CYS-RICH.

FT	DOMAIN	473	646	CYS-RICH.
FT	DOMAIN	722	989	PROTEIN KINASE.
FT	NP_BIND	728	736	ATP (BY SIMILARITY).
FT	BINDING	755	755	ATP (BY SIMILARITY).
FT	ACT_SITE	847	847	BY SIMILARITY.
FT	DISULFID	196	205	BY SIMILARITY.
FT	DISULFID	200	213	BY SIMILARITY.
FT	DISULFID	221	228	BY SIMILARITY.
FT	DISULFID	225	236	BY SIMILARITY.
FT	DISULFID	237	245	BY SIMILARITY.
FT	DISULFID	241	253	BY SIMILARITY.
FT	DISULFID	256	265	BY SIMILARITY.
FT	DISULFID	269	296	BY SIMILARITY.
FT	DISULFID	300	312	BY SIMILARITY.
FT	DISULFID	316	332	BY SIMILARITY.
FT	DISULFID	335	339	BY SIMILARITY.
FT	DISULFID	513	522	BY SIMILARITY.
FT	DISULFID	517	530	BY SIMILARITY.
FT	DISULFID	533	542	BY SIMILARITY.
FT	DISULFID	546	562	BY SIMILARITY.
FT	DISULFID	565	578	BY SIMILARITY.
FT	DISULFID	569	586	BY SIMILARITY.
FT	DISULFID	589	598	BY SIMILARITY.
FT	DISULFID	602	625	BY SIMILARITY.
FT	DISULFID	628	636	BY SIMILARITY.
FT	DISULFID	632	644	BY SIMILARITY.
FT	MOD_RES	1141	1141	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	MOD_RES	1250	1250	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	CARBOHYD	68	68	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	188	188	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	260	260	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	532	532	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	573	573	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	631	631	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	661	661	V -> E (IN ONCOGENIC NEU).
SO	SEQUENCE	1257 AA;	138831 MW;	6129264583011402 CRC64;
Query Match 70.3%; Score 1608.5; DB 1; Length 1257;				
Best Local Similarity 85.0%; Pred. No. 5.3e-117;				
Matches 294; Conservative 15; Mismatches 36; Indels 1; Gaps 1				
OY	1	MELALACRWGLLALLPEGA	STOVCTGTDMLRLPAS	PETHLMDLRHLHYGCGVOVGML 60
DB	1	MELAAWCRKGFLLALPEI	AGTVOCTGCTDMKRLRPA	SPEPHIDMLRHLYOGCGVOVGML 60
OY	61	ELITYPTNANSLFLODIO	GVGYVLIHANOVOVPLOR	BLRIYRGTOLEFDNYALAVLDNG 120
DB	61	ELITYPTNANSLFLODIO	GVGYVLIHANOVARVPLOR	BLRIYRGTOLEFDKTLAVLDNM 120
OY	121	DELNTTTPVT-GASPG	ELRELOLRSTELKSGVLI	QORNPOLCYODTILMKDIFHKNNDL 179
DB	121	DPQDNVAASPTPE	RTEGRELQRLSRSTELK	SGVLIQRNPOLCYODMVLKMDVFRKNNDL 180
OY	180	ALTLIDTNRSRACRHC	SPMKCGSRCKWGESSE	EDCCOSLTRIVACAGCARCGRLPTDCCHD 239
DB	181	AAVDIDTNNRSRACPC	APACAPKCDNHGWES	PEDCOILTGFTICSGCARCGRLPTDCCHD 240
OY	240	CAAGCTGPRHSGCLAC	LHNHNSGICELHCPAL	VLYTYTDTTFESPNPNEGRTFCGACVTCAC 299
DB	241	CAAGCTGPRHSGCLAC	LHNHNSGICELHCPAL	VLYTYTDTTFESMHNNEGRTTGACVTCAC 300
OY	300	PYNVLYSTDVGSCTLV	CPHLNDEVTAEADGT	QORCEKCSKPCARCSTHSL 345
DB	301	PYNVLYSTEVGSCTLV	CPRNQDEVTAEADGT	QORCEKCSKPCARVCYGL 346
RESULT 3				
FRB2_MESAU				
AC 060553; STANDARD; PRT; 1254 AA.				
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			

```

DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)
DE (p18erbB2) (NEU proto-oncogene) (C-erbB-2).
GN ERBB2 OR NEU.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Neve;
RX MEDLINE=94193007; PubMed=7908275;
RA Nakamura T., Ushijima T., Ishizaka Y., Nagao M., Arai M.,
RA Yamazaki Y., Ishikawa T.;
RT "Cloning and activation of the Syrian hamster neu proto-oncogene.";
RL Gene 140:251-255(1994).
CC -1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,
CC ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A
CC POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-
CC ALPHA AND AMPHIREGULIN (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS
CC (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
CC RESIDUES.
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D16295; BAA03801.1; -.
CC HSP: P11362; IFGK.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR000719; Euk_kinase.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR001245; Tyr_kinase.
DR InterPro: IPR004019; YLP_motif.
DR Pfam: PF00069; Kinase; 1.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR Pfam: PF02757; YLP; 2.
DR ProDom: PD000001; Euk_kinase; 1.
DR SMART: SM00261; FU; 3.
DR SMART: SM00219; TYKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW Proto-oncogene; Disease mutation.
FT SIGNAL 1..21 POTENTIAL.
FT CHAIN 22..1254 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.
FT DOMAIN 22..652 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 653..675 POTENTIAL.
FT DOMAIN 676..1254 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 158..368 CYS-RICH.
FT DOMAIN 472..644 CYS-RICH.
FT DOMAIN 720..987 PROTEIN KINASE.
FT NP_BIND 726..734 ATP (BY SIMILARITY).
FT BINDING 753..753 ATP (BY SIMILARITY).
FT ACT_SITE 845..845 BY SIMILARITY.
FT DISULFID 195..204 BY SIMILARITY.
FT DISULFID 199..212 BY SIMILARITY.
FT DISULFID 236..244 BY SIMILARITY.
FT DISULFID 240..252 BY SIMILARITY.
FT DISULFID 255..264 BY SIMILARITY.

```

```

FT DISULFID 268..295 BY SIMILARITY.
FT DISULFID 299..311 BY SIMILARITY.
FT DISULFID 315..331 BY SIMILARITY.
FT DISULFID 334..338 BY SIMILARITY.
FT DISULFID 344..350 BY SIMILARITY.
FT DISULFID 511..520 BY SIMILARITY.
FT DISULFID 515..528 BY SIMILARITY.
FT DISULFID 531..540 BY SIMILARITY.
FT DISULFID 544..560 BY SIMILARITY.
FT DISULFID 563..576 BY SIMILARITY.
FT DISULFID 567..584 BY SIMILARITY.
FT DISULFID 587..596 BY SIMILARITY.
FT DISULFID 600..623 BY SIMILARITY.
FT DISULFID 626..634 BY SIMILARITY.
FT DISULFID 630..642 BY SIMILARITY.
FT MOD_RES 1139..1139 BY SIMILARITY.
FT MOD_RES 1247..1247 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 68..68 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 125..125 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 187..187 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 259..259 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 530..530 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 571..571 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 629..629 N-LINKED (GLCNAC...) (POTENTIAL).
FT VARIANT 658..658 V -> E (IN ONCOGENIC NEU).
FT VARIANT 659..659 V -> E (IN ONCOGENIC NEU).
SQ SEQUENCE 1254 AA; 138252 MW; 974C3791C21F2BB1 CRC64;

```

Query Match 68.7%; Score 1571; DB 1; Length 1254;

Best Local Similarity 74.4%; Pred. No. 4, 2e-114;

Matches 299; Conservative 26; Mismatches 67; Indels 10; Gaps 3;

```

QY 1 MEALALCRWGLLALLPPGASTVCTGDMKRLRASPETHIDMLRHLYOGGVQGNL 60
DB 1 MELANCGWLLALLSPAGSGVCTGDMKRLRASPETHIDYRHLXOGGVQGNL 60
QY 61 ELTYLPNASLSFLQDIQEVQGVYLAHNOVROVPLQRLRIVRGTOLEFDNALAVLNG 120
DB 61 ELTYLPANATLSFLQDIQEVQGYMLAHNSQVRHVPRLRLIVRGTOLEFDKYALAVLNR 120
QY 121 DPLNNTPTVVGASRGGLREQLSLREILKGVLIORNPOLCYODPIIMKDFHKNNOA 180
DB 121 DPLDNNVTATGRTPEGRRELQLSLREILKGVLIORNPOLCYQDVTVMDFVRKNQDA 180
QY 181 LTLIDTNRSRACPCSPMCKSGRCWSESEDCSLTRVACAGCAGCAGKGLPTDCCHEOC 240
DB 181 PVDIDNRSRACPPACPACKDNMCWASPEDCQTLGTIAPRAVPAARARLPDCCHEOC 240
QY 241 AAGCTGKRHSDCIACLPNHSIGICELHCPALVYNTDFESMPNPGRYTFGASCVTACP 300
DB 241 AAGCTGKRHSDCIACLPNHSIGICELHCPALVYNTDFESMPNPGRYTFGASCVTACP 300
QY 301 YNLTSDVSGCTVCPPLAHNOEVAEDGTORCEKSCPKAR-----GHSLLPRAVPAV 355
DB 301 YNLTSDVSGCTVCPPLAHNOEVAEDGTORCEKSCSACVAYGLDMHRLGARALTSAN 360
QY 356 LRMPQPG--PAHPVLSFLRPSWDLVSAYSLPLAPLSPTSVPI 395
DB 361 IQEPAGCKRIFGSLAPLPRESFD---GNPSSGIAPLPEDQLQV 399

```

RESULT 4

```

ID EGRF_CHICK STANDARD: PRT: 703 AA.
AC P13387;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (CER)
DE (Fragment).
GN EGRF.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

```


RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE-Placenta;
 RX MEDLINE=97078686; PubMed=8918811;
 RA Reiter J.L., Maithe N.J.;
 RT "A 1.8 kb alternative transcript from the human epidermal growth factor receptor gene encodes a truncated form of the receptor.";
 RL Nucleic Acids Res. 24:4050-4056(1996).
 RN [14]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE-Placenta;
 RX MEDLINE=97256547; PubMed=9103388;
 RA Ilekis J.V., Garfitt J., Niederberger C., Scoccia B.;
 RT "Expression of a truncated epidermal growth factor receptor-like protein (TEGFR) in ovarian cancer";
 RL Gynecol. Oncol. 65:36-41(1997).
 RN [15]
 RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
 RC TISSUE-Placenta;
 RX MEDLINE=21100872; PubMed=11161793;
 RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J., Schehl Sinclair C., Pearsall R.S., Green P.J., Yee D., Lampland A.L., Balasubramaniam S., Crossley T.D., Magnuson T.R., James C.D., Maithe N.J.;
 RT "Comparative genomic sequence analysis and isolation of human and mouse alternative Egfr transcripts encoding truncated receptor isoforms.";
 RL Genomics 71:1-20(2001).
 RN [16]
 RP SEQUENCE OF 575-687 FROM N.A.
 RA Reiter J.L., Threadgill D.W., Danielson A.J., Schehl C.M., Lampland A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R., Maithe N.J.;
 RT "Human and mouse alternative Egfr transcripts encoding only the extracellular domain of the receptor";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [17]
 RP SEQUENCE OF 713-924 FROM N.A.
 RX MEDLINE=84196372; PubMed=6326261;
 RA Lin C.R., Chen W.S., Krueger W., Stolarsky L.S., Weber W., Evans R.M., Verma I.M., Gill G.N., Rosenfeld M.G.;
 RT "Expression cloning of human EGF receptor complementary DNA: gene amplification and three related messenger RNA products in A431 cells.";
 RL Science 224:843-848(1984).
 RN [18]
 RP SEQUENCE OF 150-962 FROM N.A.
 RX MEDLINE=84245835; PubMed=6330563;
 RA Xu Y.H., Ishii S., Clark A.J.L., Sullivan M., Wilson R.K., Ma D.P., Roe B.A., Merlino G.T., Pastan I.;
 RT "Human epidermal growth factor receptor cDNA is homologous to a variety of RNAs overproduced in A431 carcinoma cells.";
 RL Nature 309:806-810(1984).
 RN [19]
 RP SEQUENCE OF 1028-1210 FROM N.A.
 RX MEDLINE=85046463; PubMed=6093780;
 RA Simmen F.A., Goepf M.L., Schultz T.Z., Wright D.A., Carpenter G., O'Malley B.W.;
 RT "Isolation of an evolutionarily conserved epidermal growth factor receptor cDNA from human A431 carcinoma cells";
 RL Biochem. Biophys. Res. Commun. 124:125-132(1984).
 RN [110]
 RP SEQUENCE OF 1-29 FROM N.A.
 RX MEDLINE=88217333; PubMed=3329716;
 RA Haley J.D., Whitte N., Bennett P., Kinchington D., Ullrich A., Waterfield M.D.;
 RT "The human EGF receptor gene: structure of the 110 kb locus and identification of sequences regulating its transcription.";
 RL Oncogene Res. 1:375-396(1987).
 RN [111]
 RP SEQUENCE OF 1-29 FROM N.A.
 RX MEDLINE=91107677; PubMed=1988448;
 RA Haley J.D., Waterfield M.D.;
 RT "Contributory effects of de novo transcription and premature

RT transcript termination in the regulation of human epidermal growth factor receptor proto-oncogene RNA synthesis.";
 RL J. Biol. Chem. 266:1746-1753(1991).
 RN [112]
 RP SEQUENCE OF 1-29 FROM N.A.
 RX MEDLINE=85270438; PubMed=2991899;
 RA Ishii S., Xu Y.H., Stratton R.H., Roe B.A., Merlino G.T., Pastan I.;
 RT "Characterization and sequence of the promoter region of the human epidermal growth factor receptor gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:4920-4924(1985).
 RN [113]
 RP SEQUENCE OF 540.
 RA Kohda D.;
 RT Submitted (SEP-1997) to the SWISS-PROT data bank.
 RN [114]
 RP RECEPTOR ACTIVITY.
 RX MEDLINE=84191554; PubMed=6325948;
 RA Mroczkowski B., Mosig G., Cohen S.;
 RT "ATP-stimulated interaction between epidermal growth factor receptor and supercoiled DNA.";
 RL Nature 309:270-273(1984).
 RN [115]
 RP PHOSPHORYLATION.
 RX MEDLINE=89278137; PubMed=2543678;
 RA Margolis B.L., Iax I., Kris R., Dombalagian M., Honegger A.M., Howk R., Glyvol D., Ullrich A., Schlessinger J.;
 RT "All autophosphorylation sites of epidermal growth factor (EGF) receptor and HER2/neu are located in their carboxyl-terminal tails. Identification of a novel site in EGF receptor.";
 RL J. Biol. Chem. 264:10667-10671(1989).
 RN [116]
 RP CARBOHYDRATE-LINKAGE SITES ASN-128; ASN-175; ASN-413; ASN-444 AND ASN-528.
 RX MEDLINE=96398132; PubMed=8962717;
 RA Smith K.D., Davies M.J., Bailey D., Renout D.V., Hounsell E.F.;
 RT "Analysis of the glycosylation patterns of the extracellular domain of the epidermal growth factor receptor expressed in Chinese hamster ovary fibroblasts.";
 RL Growth Factors 13:121-132(1996).
 RN [117]
 RP CARBOHYDRATE-LINKAGE SITES ASN-56; ASN-352; ASN-361; ASN-568 AND ASN-603.
 RX MEDLINE=20198209; PubMed=10731668;
 RA Sato C., Kim J.-H., Abe Y., Saito K., Yokoyama S., Kohda D.;
 RT "Characterization of the N-Oligosaccharides attached to the atypical Asn-X-Cys sequence of recombinant human epidermal growth factor receptor.";
 RL J. Biochem. 127:65-72(2000).
 RN [118]
 RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
 RX MEDLINE=98225196; PubMed=9556602;
 RA Abe Y., Odaka M., Inagaki F., Iax I., Schlessinger J., Kohda D.;
 RT "Disulfide bond structure of human epidermal growth factor receptor.";
 RL J. Biol. Chem. 273:11150-11157(1998).
 RN [119]
 RP REVIEW.
 RX MEDLINE=87297456; PubMed=3039909;
 RA Carpenter G.;
 RT "Receptors for epidermal growth factor and other polypeptide mitogens.";
 RL Annu. Rev. Biochem. 56:881-914(1987).
 CC -1- FUNCTION: Receptor for EGF, but also for other members of the EGF family, as TGF-alpha, amphiregulin, betacellulin, heparin-binding EGF-like growth factor, Gp30 and vaccinia virus growth factor. It is involved in the control of cell growth and differentiation.
 CC -1- FUNCTION: Isoform 2/truncated isoform may act as an antagonist. tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Isoform 2 is secreted
 CC -1- ALTERNATIVE PRODUCTS: 4 isoforms; 1/p170 (shown here), 2/p60/ truncated isoform/TEGFR, 3/p110 and 4; are produced by alternative splicing.

DR	PROSITE	PS00107	PROTEIN_KINASE_ATP_1	
DR	PROSITE	PS00109	PROTEIN_KINASE_TPR_1	
DR	PROSITE	PS50011	PROTEIN_KINASE_DOM_1	
KW	Transmembrane; Glycoprotein; Receptor; Signal; Transferase;			
KM	Tyrosine-protein kinase; ATP-binding; Phosphorylation; Repeat.			
FT	SIGNAL	1	24	POTENTIAL.
FT	CHAIN	25	1210	POTENTIAL GROWTH FACTOR RECEPTOR.
FT	DOMAIN	25	647	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	648	670	POTENTIAL.
FT	DOMAIN	671	1210	CYTOLASMIC (POTENTIAL).
FT	REPEAT	75	300	APPROXIMATE.
FT	REPEAT	390	600	APPROXIMATE.
FT	DOMAIN	1028	1071	SER-RICH.
FT	DOMAIN	714	981	PROTEIN KINASE.
FT	NP_BIND	720	728	ATP (BY SIMILARITY).
FT	NP_BIND	747	747	ATP (BY SIMILARITY).
FT	BINDING	839	839	BY SIMILARITY.
FT	ACT_SITE	190	199	BY SIMILARITY.
FT	DISULFID	190	199	BY SIMILARITY.
FT	DISULFID	194	207	BY SIMILARITY.
FT	DISULFID	215	223	BY SIMILARITY.
FT	DISULFID	219	231	BY SIMILARITY.
FT	DISULFID	232	240	BY SIMILARITY.
FT	DISULFID	236	248	BY SIMILARITY.
FT	DISULFID	251	260	BY SIMILARITY.
FT	DISULFID	264	291	BY SIMILARITY.
FT	DISULFID	295	307	BY SIMILARITY.
FT	DISULFID	311	326	BY SIMILARITY.
FT	DISULFID	329	333	BY SIMILARITY.
FT	DISULFID	506	515	BY SIMILARITY.
FT	DISULFID	510	523	BY SIMILARITY.
FT	DISULFID	526	535	BY SIMILARITY.
FT	DISULFID	539	555	BY SIMILARITY.
FT	DISULFID	558	571	BY SIMILARITY.
FT	DISULFID	562	579	BY SIMILARITY.
FT	DISULFID	582	591	BY SIMILARITY.
FT	DISULFID	595	617	BY SIMILARITY.
FT	DISULFID	620	628	BY SIMILARITY.
FT	DISULFID	624	636	BY SIMILARITY.
FT	MOD_RES	680	680	PHOSPHORYLATION (BY PKC).
FT	MOD_RES	1092	1092	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	MOD_RES	1110	1110	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	MOD_RES	1172	1172	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	MOD_RES	1197	1197	PHOSPHORYLATION (AUTO-, MAJOR SITE) (BY SIMILARITY).
FT	CARBOHYD	128	128	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	175	175	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	196	196	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	352	352	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	413	413	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	444	444	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	528	528	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	568	568	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	603	603	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	623	623	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	19	19	C -> S (IN REF. 2).
FT	CONFLICT	539	539	C -> W (IN REF. 5).
FT	CONFLICT	991	991	L -> F (IN REF. 4).
FT	CONFLICT	1116	1117	HP -> DR (IN REF. 6).
FO	SEQUENCE	1210	AA: 134853	MM: 6590E20DA6FD2D2F5 CRC64;

	Query Match	34.5%	Score 789;	DB 1;	Length 1210;	
	Best Local Similarity	46.3%;	Pseq.No.	1.7e-53;		
	Matches 156; Conservative	43;	Msmatches	116;	Indels 22;	Gaps 6
OY	11 LLLALLPPCAA--STQVCGTDMKRLRASPETHMLRHLYOGGVOCVOGNELTYLPTN	68				
	: :	:	: : :: :	:		
Dd	14 LTATLCAMGGALEEKKVCQGTSNRILQTGFEDHFLSLDPMYNCEVVGLNGLEITYVRN	73				
OY	69 ASLSLELDIDEOVGYYVLINHNQVRQLRIIVRGTFOLFDEVDYAVALDNCPPLNNTPTR	128				
	:: :	:	: : :: :	:		
Dd	74 YDLSTLKTIQEAVGYVLIALNTVERPLENDLIQRNALXYENTAYALAIISN-----	124				
OY	139 VTGASPGGLEIRQLNSLTIELKGVIHQENPOLCYDTITIMDI---FHKNQAALTII	184				

```

Db      125  -YGRNTRGRLRELPRKRLQETILLICAVFRESNNPILCNNDTIQMRDIYQNVFMMSMDL- 180
Oy      185  DTNRSRACHPDCSPKCKSGRCWCESSSEPDQSLRTVYCAGCA-RCKGPLPTDCCBQCAAG 243
Db      181  -QSHPSQCPKDCPSCPGSGCWGGGEENCKLFRILICAQCSHRRCGRBSPSDCHNCAAG 239
Oy      244  CTGRKHSQCLACLAHFNINSIGCELCALCALTYNTNTDTFESMPNREGRTGASCVTAQCPNY 303
Db      240  CTGRESCELCVCKCFQDEATCKDCTCPMLMLNPTTYQMVNVEGKYSFGATCVKCKPKNY 299
Oy      304  LSTDVGSCTLWCPLHNOEYVAEDGTORCEKCSRPCAR 340
Db      300  VYTHGSCVRACGPDYIEV-EDDGIRCKKCDGPKRK 335

RESULT 7
ERB4_RAT
ID      ERB4_RAT      STANDARD:      PRT: 1308 AA.
AC      Q62936; Q922N7;
DT      15-DEC-1998 (Rel. 37, Created)
DT      15-JUN-2002 (Rel. 41, Last sequence update)
DE      15-JUN-2002 (Rel. 41, Last annotation update)
DT      Receptor protein-tyrosine kinase erbB-4 precursor (EC 2.7.1.112).
GN      ERB4 OR TYRO-2.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_Taxid=101116;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Heart;
RA      MEDLINE=98221155; PubMed=9553078;
RA      Zhao Y.-Y., Sawyer D.R., Balliga R.R., Opel D.J., Han X.,
RA      Marchionni M.A., Kelly R.A.;
RT      "Neuregulins promote survival and growth of cardiac myocytes.
RT      Persistence of ErbB2 and ErbB4 expression in neonatal and adult
RT      ventricular myocytes.";
RL      J. Biol. Chem. 273:10261-10269(1998).
RN      [12]
RP      SEQUENCE OF 848-901 FROM N.A.
RC      TISSUE=Sciatic nerve;
RX      MEDLINE=9122560; PubMed=2025425;
RA      Lai C., Lemke G.;
RT      "An extended family of protein-tyrosine kinase genes differentially
RT      expressed in the vertebrate nervous system.";
RL      Neuron 6:691-704(1991).
RN      [13]
RP      SEQUENCE OF 1031-1198 FROM N.A.
RC      STRAIN=Sprague-Dawley; TISSUE=spinal cord;
RX      MEDLINE=97184212; PubMed=9030624;
RA      Carroll S.L., Miller M.L., Frohnert P.W., Kim S.S., Corbett J.A.;
RT      "Expression of neuregulins and their putative receptors, ErbB2 and
RT      ErbB3, is induced during Wallerian degeneration.";
RL      J. Neurosci. 17:1642-1659(1997).
RN      [1]
RP      FUNCTION: SPECIFICALLY BINDS AND IS ACTIVATED BY NEUREGULINS, NRG-
RN      2, NRG-3, HEPARIN-BINDING EGF-LIKE GROWTH FACTOR, BETACELLULIN AND
RN      NNAK. INTERACTION WITH THESE FACTORS INDUCES CELL DIFFERENTIATION,
RN      NOT ACTIVATED BY EGF, TGF- $\alpha$ , AND AMPHIREGULIN (BY SIMILARITY).
CC      CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein
CC      tyrosine phosphate.
CC      -1- SUBUNIT: HOMODIMER OR HETERODIMER WITH EACH OF THE OTHER ERBB
CC      RECEPTORS (POTENTIAL).
CC      -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC      -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN THE DEVELOPING
CC      NERVOUS SYSTEM. EXHIBITS DISTINCT AND HIGHLY REGIONALIZED PATTERNS
CC      OF EXPRESSION IN THE ADULT BRAIN, WHERE IT IS MAINLY FOUND IN THE
CC      RETICULAR NUCLEUS OF THE THALAMUS. VERY LOW LEVELS IN KIDNEY, AND
CC      HEART.
CC      -1- PMW: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
CC      RESIDUES (BY SIMILARITY).
CC      -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.

```


RT "Isolation and characterization of ERBB3, a third member of the
 RT ERBB/epidermal growth factor receptor family: evidence for
 RT overexpression in a subset of human mammary tumors."
 RL Proc. Natl. Acad. Sci. U.S.A. 86:9193-9197(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90311312; PubMed=2164210;
 RA Plowman G.D., Whitney G.S., Neubauer M.G., Green J.M., McDonald V.L.,
 RA Todaro G.J., Shoyab M.;
 RT "Molecular cloning and expression of an additional epidermal growth
 RT factor receptor-related gene."
 RL Proc. Natl. Acad. Sci. U.S.A. 87:4905-4909(1990).
 RN [3]
 RP SEQUENCE FROM N.A. (SHORT FORM).
 RC TISSUE=Placenta;
 RX MEDLINE=93282822; PubMed=7685162;
 RA Katoh M., Yazaki Y., Sugimura T., Terada M.;
 RT "c-erbB gene encodes secreted as well as transmembrane receptor
 RT tyrosine kinase."
 CC Biochem. Biophys. Res. Commun. 192:1189-1197(1993).
 CC -1- FUNCTION: BINDS AND IS ACTIVATED BY NEUREGULIN AND NTAK.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS
 CC (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (LONG FORM) AND
 CC SECRETED (SHORT FORM).
 CC -1- ALTERNATIVE PRODUCTS: TWO FORMS, A LONG TRANSMEMBRANE FORM AND A
 CC SHORT SECRETED FORM. EXIST DUE TO ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EPITHELIAL TISSUES AND BRAIN.
 CC -1- DOMAIN: THE CYTOPLASMIC PART OF THE RECEPTOR MAY INTERACT WITH THE
 CC SH2 OR SH3 DOMAINS OF MANY SIGNAL-TRANSDUCING PROTEINS.
 CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES
 CC AND PROMOTES ITS ASSOCIATION WITH THE p65 SUBUNIT OF
 CC PHOSPHATIDYLINOSITOL 3-KINASE (BY SIMILARITY).
 CC -1- DISEASE: OVEREXPRESSED IN A SUBSET OF HUMAN MAMMARY TUMORS.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL; M29366; AAA35790.1; -
 DR EMBL; M34309; AAA35979.1; -
 DR EMBL; S61953; AAB26935.1; -
 DR PIR; A36223; A36223.
 DR HSSP; P11362; 1FGK.
 DR GeneW; HGNC:3431; ERBB3.
 DR MIM; 190151; -
 DR InterPro; IPR000494; EGFR_L_domain.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR002174; Furin-like.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase.1.
 DR Pfam; PF00757; Furin-like.1.
 DR Pfam; PF01030; Recep_L_domain.2.
 DR ProDom; PD000001; Euk_pkinase.1.
 DR SMART; SM00261; FU; 3.
 DR SMART; SMO0219; TyrKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; FALSE_NEG.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
 KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 KW Alternative splicing.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 1342 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-3.
 FT DOMAIN 20 643 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 644 664 POTENTIAL.

FT	DOMAIN	665	1342	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	709	966	PROTEIN KINASE.
FT	NP_BIND	715	723	ATP (BY SIMILARITY).
FT	BINDING	742	742	ATP (BY SIMILARITY).
FT	ACT_SITE	834	834	BY SIMILARITY.
FT	DISULFID	186	194	BY SIMILARITY.
FT	DISULFID	190	202	BY SIMILARITY.
FT	DISULFID	210	218	BY SIMILARITY.
FT	DISULFID	214	226	BY SIMILARITY.
FT	DISULFID	227	235	BY SIMILARITY.
FT	DISULFID	231	243	BY SIMILARITY.
FT	DISULFID	246	255	BY SIMILARITY.
FT	DISULFID	259	286	BY SIMILARITY.
FT	DISULFID	290	301	BY SIMILARITY.
FT	DISULFID	305	320	BY SIMILARITY.
FT	DISULFID	323	327	BY SIMILARITY.
FT	DISULFID	500	509	BY SIMILARITY.
FT	DISULFID	504	517	BY SIMILARITY.
FT	DISULFID	520	529	BY SIMILARITY.
FT	DISULFID	533	549	BY SIMILARITY.
FT	DISULFID	552	565	BY SIMILARITY.
FT	DISULFID	556	573	BY SIMILARITY.
FT	DISULFID	576	585	BY SIMILARITY.
FT	DISULFID	589	610	BY SIMILARITY.
FT	DISULFID	613	621	BY SIMILARITY.
FT	DISULFID	617	629	BY SIMILARITY.
FT	CARBOHYD	126	126	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	250	250	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	353	353	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	408	408	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	414	414	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	437	437	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	469	469	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	522	522	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	566	566	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	616	616	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	VARSPLIC	141	183	ELISGVYIEKNDKLCHEMDTIDMDIVDRDAEIVKDNGR SC -> GQFVPSGLTFPOGQDQWLLDDDRRLTLSSASK VPTVLAIV (IN SHORT ISOFORM). MISSING (IN SHORT ISOFORM). E -> G (IN REF. 2). E -> G (IN REF. 2).
FT	VARSPLIC	184	1342	
FT	CONFLICT	560	560	
FT	CONFLICT	1064	1064	
SO	SEQUENCE	1342 AA;	148097 MW;	7201E766CA374BD CRC64;

Query Match 32.2%; Score 735.5; DB 1; Length 1342;
 Best Local Similarity 44.0%; Pred. No. 2; 7e-49;
 Matches 150; Conservative 49; Mismatches 121; Indels 21; Gaps 9;

QY	10	GILLALLPPGAA--STOVCTGTDMKRLRPASPEITLDMRLHYGCGOVVGNLEIYLP	67
DB	11	GLLFSLARSEVGNQAVCPGTLNLGSLVYGDABENQYOTILKLYERCEVYVMGLIEVLVGH	70
QY	68	NASLSFLDIOEVGCVLIANQVROVPLQRLIRIVRGTLQFEDNYALAVLDNGDPLNNTT	127
DB	71	NADLSFLQWIREVGYVIVAMNEFSTLPLNLRVVRGQVYDGKRAIFVM-----LNNWT	125
QY	128	PYTGASPGGLRLOLRSLTEILKGVLLQORNPOLCYODTILMKDIFKNNQALLLIDTN	187
DB	126	---NSSHALRQLRLTQLEILLGSGVYIEKNDKLCHEMDTIDMDIVDRD-----AEIVAKD	178
QY	188	RSRACHPSCPMCKSGRCSESSDCOSLTRVACAGGC--ARCGRLPTPOCHQCAAGCG	246
DB	179	NKRSCLPCHVEYKRG--KCMGPGSEDCOTLTKTICAPQCNCHCGPMPNOCCHDECAAGCGSG	237
QY	247	PKHSDCLACLFHNHSGICELNCPALVYNTDTFESMPNDEGRYTGASCAYTACPNYLS	306
DB	238	PODTCFACRHNFDNSGACVPCRPQPLVYNNKLTFLQLEPMPHRYQYGVGVASCAPHNFV--V	296
QY	307	DVSGCTIVCPNLNQEVTAEDEGRQREKSKPCAR---GTHS	344
DB	297	DOTSCVRACPPDKMEVD--KNGIKMCEPGGLCPKACEGTGS	336

```
RESULT 10
ERB3_RAT STANDARD: PRT: 1339 AA.
AC 062799; Q62955;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-3 precursor (EC 2.7.1.112)
DE (c-erbB3).
GN ERBB3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=96096535; PubMed=8522190;
RA Hellyer N.J., Kim H.-H., Greaves C.H., Sierke S.L., Koland J.G.;
RT "Cloning of the rat ErbB3 cDNA and characterization of the
RT recombinant protein."
RT Gene 165:279-284(1995).
RN (2)
RP REVISIONS TO 85: 513 AND 565.
RA Hellyer N.J., Koland J.G.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN (3)
RP SEQUENCE OF 922-1097 FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE=Sciatic nerve;
RX MEDLINE=97184212; PubMed=9030624;
RA Carroll S.L., Miller M.L., Frohner P.W., Kim S.S., Corbett J.A.;
RT "Expression of neuroligins and their putative receptors, ErbB2 and
RT ErbB3, is induced during Wallerian degeneration."
RT J. Neurosci. 17:1642-1659(1997).
CC -I- FUNCTION: BINDS AND IS ACTIVATED BY NEUREGULINS AND NTAK.
CC -I- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -I- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS
CC (POTENTIAL).
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
CC -I- DOMAIN: THE CYTOPLASMIC PART OF THE RECEPTOR MAY INTERACT WITH THE
CC SH2 OR SH3 DOMAINS OF MANY SIGNAL-TRANSDUCING PROTEINS.
CC -I- PPM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES
CC AND PROMOTES ITS ASSOCIATION WITH THE P85 SUBUNIT OF
CC PHOSPHATIDYLINOSITOL 3-KINASE.
CC -I- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC -----
CC This SWISS-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch)
CC -----
DR EMBL: U29339; AAC28498.2; -
DR EMBL: U52530; AAC53050.1; -
DR HSSP: P11362; IFKC.
DR InterPro: IPR000494; EGF_R_L_domain.
DR InterPro: IPR000719; Euk_kinase.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR001245; TYR_kinase.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PD000001; Euk_kinase; 1.
DR SMART: SM00261; FU; 5.
DR SMART: SM00219; TYRK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
```

```
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT SIGNAL 1 19
FT CHAIN 20 1339
FT DOMAIN 20 643
FT TRANSFM 644 662
FT DOMAIN 663 1339
FT DOMAIN 183 259
FT DOMAIN 707 964
FT NP_BIND 713 721
FT BINDING 740 740
FT ACT_SITE 832 832
FT DISULFD 186 194
FT DISULFD 190 202
FT DISULFD 210 218
FT DISULFD 214 226
FT DISULFD 227 235
FT DISULFD 231 243
FT DISULFD 246 255
FT DISULFD 259 286
FT DISULFD 290 301
FT DISULFD 305 320
FT DISULFD 323 327
FT DISULFD 500 509
FT DISULFD 504 517
FT DISULFD 520 529
FT DISULFD 533 549
FT DISULFD 556 573
FT DISULFD 576 585
FT DISULFD 589 610
FT DISULFD 613 621
FT DISULFD 617 629
FT CARBOHYD 126 126
FT CARBOHYD 250 250
FT CARBOHYD 353 353
FT CARBOHYD 408 408
FT CARBOHYD 414 414
FT CARBOHYD 437 437
FT CARBOHYD 469 469
FT CARBOHYD 522 522
FT CARBOHYD 566 566
FT CARBOHYD 616 616
FT CONFLICT 1028 1028
SQ SEQUENCE 1339 AA; 147545 MW; 0AA5F2402BBFD1E CRC64;
Query Match 30.5%; Score 698; DB 1; Length 1339;
Best Local Similarity 42.1%; Pred. No. 2,2e-46;
Matches 147; Conservative 46; Mismatches 130; Indels 26; Gaps 10;
```

QY	3	LAALCRWGLLLALPPGAA---STOYCTGTDMKRLPASPETHLDMRLHYGGCOVQGN	59
DB	7	LQVLC---FLSLARGSEMGNSQAVCPDTLNGLSVTGDADNOYOTLYKYLERKECEVMGN	62
QY	60	LEFLYLPNTASLPFODIOEVGYVLIAHNOVROYPLQRLRVRTGOLFEDNRYALAVLDN	119
DB	63	LEIVLTGNHADSFLQWIREVTGYVLVAMNESVPLPLKRVYRGTYOYDGKALFYM--	120
QY	120	GDPLNNTTPVTGASPGGLREIQRLSTELKGVLIQRPOLCYODTLIMDKIFKNNQL	179
DB	121	---LWVNT-----NSSHALRLQKFTQLETILSGVYIERKDKLCHMDTIDMRIVR--	170
QY	180	ALTLLDTNRSRACHCSPCKSRCKSGWESSEPCQSLTRIVYACGGC-ARCKGFLPTDCHE	238
DB	171	GAEIYVKNNGANCPCPCHEVCKG-RCWGPGPDQILITKIKAPQCGNGRCFGPNPQCCHD	229
QY	229	OCAAGCTGPKKSHSDCLACHFNHSGICELHCPALVYNTDFEPMNPREGRYFGASCVTA	298
DB	230	ECAGGCGSPQDITDCFRACRFNDSGACVPRCPPLVYNNKITLFQLEPNPHITKYIGGVCAVS	289
QY	299	CPYNYLSTDVGSCTLVCPPLAHNOEVTAEQTORCEKSCPKCAR---GTHS 344	
DB	290	CPHNEY-VDQTCVYACPPDKMEVD-KHGLKMCPCGCGGLCPACGGTGS 336	

```

RESULT 11
XMWK_XIPMA STANDARD: PRT: 1167 AA.
AC P13388:
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Melanoma receptor protein-tyrosine kinase precursor (EC 2.7.1.112).
GN XMWK OR TV.
OS Xiphophorus maculatus (Southern platyfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Cyprinodontiformes; Poeciliidae; Xiphophorus.
OX NCBI_TaxID=8083;
RP SEQUENCE FROM N.A.
RX MEDLINE=90015140; PubMed=2797166;
RA Wiltbrodt J., Adam D., Maltischek B., Mauelel W., Raulf F.,
RA Telling A., Robertson S.M., Scharlt M.;
RT "Novel putative receptor tyrosine kinase encoded by the melanoma-
RL inducing Ty locus in Xiphophorus.";
RL Nature 341:415-421(1989).
RN [2]
RP REVISION TO 515.
RA Scharlt M.;
RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
CC - FUNCTION: PROBABLE RECEPTOR WITH PROTEIN-TYROSINE KINASE ACTIVITY.
CC - CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC - SUBCELLULAR LOCATION: Type I membrane protein.
CC - DISEASE: INVOLVED IN PIGMENT CELLS MALIGNANT MELANOMAS.
CC - SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL, X16891; CAA34770.2; -.
DR PIR: S06142; S06142.
DR HSSP: P11362; 1FGK.
DR InterPro: IPR000494; EGFR_L.domain.
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR001290; Ser_thr_Pkinase.
DR InterPro: IPR001245; Tyr_Pkinase.
DR Pfam: PF00069; Pkinase.1.
DR Pfam: PF00757; Furin-like.1.
DR Pfam: PF01030; Recep_L.domain.2.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_Pkinase.1.
DR SMART: SM00261; FU; 5.
DR SMART: SM00220; S.TKc; 1.
DR SMART: SM00219; Tyrc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR Transmembrane; Glycoprotein; Receptor; Signal; Transferase;
KW Tyrosine-protein kinase; ATP-binding; Phosphorylation; Proto-oncogene.
KV Tyrosine-protein kinase; ATP-binding; Phosphorylation; Proto-oncogene.
FT SIGNAL 1 25
FT CHAIN 26 1167 MELANOMA RECEPTOR PROTEIN-TYROSINE
FT KINASE.
FT DOMAIN 26 642 EXTRACELLULAR (POTENTIAL).
FT TRASNMEM 643 665 POTENTIAL.
FT DOMAIN 666 1167 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 710 977 PROTEIN KINASE.
FT NP_BIND 716 724 ATP (BY SIMILARITY).
FT BINDING 743 743 ATP (BY SIMILARITY).
FT ACT_SITE 835 835 BY SIMILARITY.

```

```

FT DISULFID 195 204 BY SIMILARITY.
FT DISULFID 199 212 BY SIMILARITY.
FT DISULFID 220 228 BY SIMILARITY.
FT DISULFID 224 236 BY SIMILARITY.
FT DISULFID 237 245 BY SIMILARITY.
FT DISULFID 241 253 BY SIMILARITY.
FT DISULFID 256 265 BY SIMILARITY.
FT DISULFID 269 296 BY SIMILARITY.
FT DISULFID 300 311 BY SIMILARITY.
FT DISULFID 315 330 BY SIMILARITY.
FT DISULFID 333 337 BY SIMILARITY.
FT DISULFID 504 513 BY SIMILARITY.
FT DISULFID 508 521 BY SIMILARITY.
FT DISULFID 524 533 BY SIMILARITY.
FT DISULFID 537 553 BY SIMILARITY.
FT DISULFID 536 569 BY SIMILARITY.
FT DISULFID 560 577 BY SIMILARITY.
FT DISULFID 593 615 BY SIMILARITY.
FT DISULFID 618 626 BY SIMILARITY.
FT DISULFID 622 634 BY SIMILARITY.
FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 365 365 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 417 417 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 501 501 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 576 576 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1167 AA; 129934 MW; 4793E0749DCID55A CRC64;

```

Query Match 29.8%; Score 681.5; DB 1; Length 1167;

Best Local Similarity 42.4%; Pred. No. 3; Se 45;

Matches 145; Conservative 46; Mismatches 136; Indels 15; Gaps 8;

```

QY 4 AALCPGMLLALLPPEAST---QVCTGDMKRLPASPEFLDMLRHLYGCGVVOGN 59
DB 8 AALLQ--LLVLVLSISCCSTDPKRCQGSNMTM---LDNHYLKMKKMYSQCNVLE 62
QY 60 LEITYLPNMSLSEFLDDIOEVGVLIANQYRQVPLRLIRYRGTLFEDNYALAVLN 119
DB 63 LEITYQENODLSFLQSIQEVGVLIANNEVSTIPLVNLRLIRGONLYEGNETLLVMSN 122
QY 120 GDLNNTTPTVTSAPGCLAEQLRSTELLKGVLLQRPOLCYOTFIIMKDFHKNOL 179
DB 123 YOK-NPSSP--DVYQVGLKQLDLSNLTLLSGVKVSHNPLLCNVTETIMWMDIVDTSNP 179
QY 180 ALTLIDTNSRACHPGSPCKSGRCWGESSEDOSLTRFYCAGGC-ARCKGLPTDCHE 238
DB 180 TMLNLIHAFERQCKDKDHGCVMSGCAAPRGHCQKTKLKAEQCKRRRGKRPIDCNE 239
QY 239 QCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNDEGRYTGACSVTA 298
DB 240 HCAGCGTGPRAIDCLACRFENDDGTCKDPPKIDIVSHQVNDPNIKYFGACVKE 299
QY 299 CPYNYISTDVSGSTLVCPRLHNEVYTAEDGTQRCCKSKFCAR 340
DB 300 CPSNVYVTE-GACVRSASGMEVD-ENGRSKRCPCDGVCPK 339

```

```

RESULT 12
EGFR_DROME STANDARD: PRT: 1426 AA.
AC P04412; 061601; Q9W2G0; P81868;
DT 13-AUG-1987 (Rel. 05, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (Egfr)
DE (Gurken receptor) (Torpedo protein) (Drosophila relative of ERBB).
GN EGFR OR TOP OR C-ERBB OR DER OR CG10079.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;

```


OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxId=7227;
 [1]
 RP SEQUENCE FROM N.A. (ISOFORMS TYPES I AND II).
 RX MEDLINE=94350209; PubMed=8070664;
 RA Clifford R., Schupbach T.;
 RT "Molecular analysis of the *Drosophila* EGF receptor homolog reveals
 RT subdomains of the receptor protein.";
 RL Genetics 137:531-550(1994).
 [2]
 RP REVISIONS.
 RA Clifford R., Schupbach T.;
 RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RX MEDLINE=85124611; PubMed=2982499;
 RA Livneh E., Glazer L., Segal D., Schlessinger J., Shilo B.-Z.;
 RT "The *Drosophila* EGF receptor gene homolog: conservation of both
 RT hormone binding and kinase domains.";
 RL Cell 40:599-607(1985).
 [4]
 RP SEQUENCE FROM N.A., CHARACTERIZATION, AND ALTERNATIVE SPLICING.
 RC STRAIN=Oregon-R; TISSUE=Embryo;
 RX MEDLINE=87002474; PubMed=3093080;
 RA Schejter E.D., Segal D., Glazer L., Shilo B.-Z.;
 RT "Alternative 5' exons and tissue-specific expression of the
 RT *Drosophila* EGF receptor homolog transcripts.";
 RL Cell 46:1091-1101(1986).
 [5]
 RP SEQUENCE FROM N.A. (ISOFORM TYPE I), TISSUE SPECIFICITY, AND MUTATION
 RP ANALYSIS.
 RX MEDLINE=99102120; PubMed=9882502;
 RA Lesokhin A.M., Yu S.-Y., Katz J., Baker N.E.;
 RT "Several levels of EGF receptor signaling during photoreceptor
 RT specification in wild-type, Ellipse, and null mutant *Drosophila*.";
 RL Dev. Biol. 205:139-144(1999).
 [6]
 RP SEQUENCE FROM N.A. (ISOFORM TYPE I).
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceuliker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.G., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Chame M., Pfeiffer B.D.,
 RA Man K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Adganyi A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrelera S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck C.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitai B.E., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merulov G., Milshina N.V., Modarri C., Morris S., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazozo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Relbert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kimms I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 [7]
 RP SEQUENCE OF 959-1078 FROM N.A.
 RC STRAIN=Daekwanryeong;
 RX MEDLINE=85137938; PubMed=2983232;
 RA Wadsworth S.C., Vincent W.S. III, Blodeau-Wentworth D.;
 RT "A *Drosophila* genomic sequence with homology to human epidermal
 RT growth factor receptor.";
 RL Nature 314:178-180(1985).
 [8]
 RP SEQUENCE OF 1133-1137; 1155-1167 AND 1209-1216 FROM N.A., AND MUTATION
 RP ANALYSIS.
 RX MEDLINE=92038942; PubMed=1936959;
 RA Raz E., Schejter E.D., Shilo B.-Z.;
 RT "Interallelic complementation among DER/Elb alleles: implications for
 RT the mechanism of signal transduction by receptor-tyrosine kinases.";
 RL Genetics 129:191-201(1991).
 [9]
 RP REVIEW.
 RX MEDLINE=97248481; PubMed=9094709;
 RA Perrimon N., Perkins L.A.;
 RT "There must be 50 ways to rule the signal: the case of the *Drosophila*
 RT EGF receptor.";
 RL Cell 89:13-16(1997).
 CC -I- FUNCTION: BINDS TO FOUR LIGANDS: SPITZ, GURKEN, VEIN AND ARGOS,
 CC WHICH IS AN ANTAGONIST. TRANSDUCES THE SIGNAL THROUGH THE RAS-RAF-
 CC MAPK PATHWAY. INVOLVED IN A MYRIAD OF DEVELOPMENTAL DECISIONS.
 CC CRITICAL FOR THE PROLIFERATION OF IMAGINAL TISSUES. AND FOR THE
 CC DETERMINATION OF BOTH THE ANTERO-POSTERIOR AND DORSO-VENTRAL
 CC POLARITIES OF THE OOCYTE. IN THE EMBRYO, PLAYS A ROLE IN THE
 CC ESTABLISHMENT OF VENTRAL CELL FATES, MAINTENANCE OF AMNIOTEROSA
 CC AND VENTRAL NEUROECODERMAL CELLS. GERM BAND RETRACTION, CELL FATE
 CC SPECIFICATION IN THE CENTRAL NERVOUS SYSTEM AND PRODUCTION OF
 CC CUTICLE.
 CC -I- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 CC tyrosine phosphate.
 CC -I- SUBCELLULAR LOCATION: ISOFORMS TYPE I AND TYPE II ARE TYPE I
 CC MEMBRANE PROTEINS. ISOFORM TYPE III IS A TYPE III MEMBRANE
 CC PROTEIN.
 CC -I- ALTERNATIVE PRODUCTS: 3 ISOFORMS: TYPE I (SHOWN HERE), TYPE II AND
 CC TYPE III: ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -I- TISSUE SPECIFICITY: UBQUITOUSLY EXPRESSED IN EMBRYOS. IN LARVAE,
 CC UNIFORM EXPRESSION IS SEEN IN WING DISKS, GENITAL DISK, ANLAGEN OF
 CC TESTIS AND OVARY, AND BRAIN CORTEX. IN EYE-ANTENNA DISK, HIGHEST
 CC EXPRESSION IS ANTERIOR TO MORPHOGENETIC FURROW, LEVELS REMAIN HIGH
 CC IN PHOTORECEPTOR PRECURSOR CELLS. THIS PATTERN IS REVERSED IN
 CC POSTERIOR EYE DISK. IN ADULTS EXPRESSION IS HIGH IN BRAIN CORTEX
 CC AND THORACIC AND ABDOMINAL GANGLIA.
 CC -I- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF052754; AAC08536.1; -
 DR EMBL: AF052753; AAC08536.1; JOINED.
 DR EMBL: AF052754; AAC08535.1; -
 DR EMBL: AF052752; AAC08535.1; JOINED.
 DR EMBL: K03054; AAA51462.1; -
 DR EMBL: K03417; AAA51460.1; -
 DR EMBL: K03416; AAA50965.1; -
 DR EMBL: K03418; AAA51461.1; -
 DR EMBL: AF109077; AAD26134.1; -


```

DR EMBL: AF109078; AAD26132.1; -.
DR EMBL: AF109082; AAD26132.1; JOINED.
DR EMBL: AF109078; AAD26133.1; -.
DR EMBL: AF109084; AAD26133.1; JOINED.
DR EMBL: AF109079; AAD26130.1; -.
DR EMBL: AF109081; AAD26130.1; JOINED.
DR EMBL: AF109079; AAD26131.1; -.
DR EMBL: AF109083; AAD26131.1; JOINED.
DR EMBL: AF109080; AAD26135.1; -.
DR EMBL: AF003454; AAF46732.1; -.
DR EMBL: X02293; CAA26157.1; -.
DR EMBL: X78920; CAA55523.1; -.
DR EMBL: X78918; CAA55521.1; -.
DR EMBL: X78919; CAA55522.1; -.
DR PIR: A00640; GOEPE.
DR HSSP: P11362; 1EGK.
DR Flybase: FBgn0003731; Egfr.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR000719; Euk_kinase.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR001245; Tyr_kinase.
DR Pfam: PF00069; kinase.1.
DR Pfam: PF00757; Furin-like.1.
DR Pfam: PF01030; Recep_L_domain.2.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_kinase.1.
DR SMART: SM00261; FU; 7.
DR SMART: SM00219; TYRK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR Transmembrane; Glycoprotein; Receptor; 1.
DR Tyrosine-protein kinase; ATP-binding; Signal; Alternative splicing;
KW developmental protein.
KM
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 1426 EPIDERMAL GROWTH FACTOR RECEPTOR.
FT DOMAIN 31 868 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 869 889 POTENTIAL.
FT DOMAIN 890 1426 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 938 1198 PROTEIN KINASE.
FT NP_BIND 944 952 ATP (BY SIMILARITY).
FT BINDING 971 971 ATP (BY SIMILARITY).
FT ACT_SITE 1063 1063 BY SIMILARITY.
FT MOD_RES 902 902 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).

```

```

Query Match 25.1%; Score 574.5; DB 1; Length 1426;
Best Local Similarity 36.3%; Pred. No. 8.8e-37;
Matches 118; Conservative 46; Mismatches 130; Indels 31; Gaps 7;

```

```

QY 24 QVCTGDMKRLPASPETHLMDLRLHYOGGVQGNLELYLPT--NALSFLDIOGVG 82
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 100 KICIGTSRLSVSPSKENHNRNLDRRTNCTYYDGNKLTWLPENLDLFLDNIRVYG 139
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 83 YVLIANQVROVPLQRLIRVGTOLF-----EDNYALAVLDNGDPLNNTTPTVGASPGGL 137
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 160 YIILSHDVKKVYFVKQILIRGRLTFLSVSEERKALFV-----TYSKM 203
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 138 RELQRLSTFLILKGVLIQNPOLCYODTLMDIFIKNNQALTLIDTKRSRACHCSP 197
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 204 YTEILPDLRLVNLGVGFHNHNNYIMCHMRTIQMSEIYSNGDAYYNDFTAPERCCPCHE 263
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 198 MCKGRMCWSSSECCSLFTVCAGCA--RCRGPLPTDCCHCCAGCGPKHSDIAC 255
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 264 SCRHG--CWGGRPNCKCFSLTSCPOCAGRCYGPFRRECHLFCAGCGTGPPOKDIAC 322
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 256 LAFNHSGICELHCPALVYNTDTFESMPNDEGRYTFGASCVTACPNYVLTDSGCTLVC 315
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 323 KNFEDEAVSKEECPMKRYNPTFTYLETNPEGKAYGATCVKECP--CHLLRDNGACVRSVC 381
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 316 PLHNGEYTAEDGTORCKCKPCAR 340
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 382 PODMKDGE-----CVPCNGPCPK 401
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

RESULT 13
LET23_CAEL
ID LET23_CAEL STANDARD; PRT; 1323 AA.
AC P24348;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE let-23 receptor protein-tyrosine kinase precursor (EC 2.7.1.112).
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=91080919; PubMed=1979659;
RX Aroian R.V., Koga M., Mendel J.E., Ohshima Y., Sternberg P.W.;
RT "The let-23 gene necessary for Caenorhabditis elegans vulval
RT induction encodes a tyrosine kinase of the EGF receptor subfamily.";
RL Nature 348:693-699(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Bristol N2; PubMed=8604137;
RA Sakai T., Koga M., Ohshima Y.;
RT "Genomic structure and 5' regulatory regions of the let-23 gene in
RT the nematode C. elegans.";
RL J. Mol. Biol. 256:548-555(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=Bristol N2;
RA Thomas K.;
RT Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RL [4]
RP MUTANTS.
RX MEDLINE=94147981; PubMed=8313880;
RA Aroian R.V., Les G.M., Sternberg P.W.;
RT "Mutations in the Caenorhabditis elegans let-23 EGFR-like gene define
RT elements important for cell-type specificity and function.";
RL EMBO J. 13:360-366(1994).
RN [5]
RP SUBCELLULAR LOCATION.
RX STRAIN=Bristol N2;
RX MEDLINE=99287744; PubMed=10359617;
RA Whitfield C.W., Bernard C., Barnes T., Hekimi S., Kim S.K.;
RT "Basolateral localization of the Caenorhabditis elegans epidermal
RT growth factor receptor in epithelial cells by the PDZ protein
RT LIN-10.";
RL Mol. Biol. Cell 10:2087-2100(1999).
CC -!- FUNCTION: Tyrosine kinase receptor required for the induction of
CC vulval differentiation. Possible receptor for the inductive signal
CC required for vulval development. Activated by lin-3 and acts by
CC way of let-60 RAS. The lin-3/let-23 pair is a simplified version
CC of the mammalian neuroligin-ERBB network.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Basolateral and
CC apical membrane of cell junctions in epithelial vulval precursor
CC cells.
CC -!- TISSUE SPECIFICITY: Vulval precursor cells.
CC -!- DEVELOPMENTAL STAGE: Expressed during L2 and L3 larval stages.
CC -!- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL: X57767; CAA40919.1; ALT_SEQ.

```


DR HSSP: P06213; 11RK.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR003962; FNIII_repeat.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR002011; RTK_kinase1.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00041; fn3; 3.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR PRINTS: PR00014; FNTPETII.
 DR PRINTS: PR00109; TYRKINASE.
 DR Prodom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00060; FN3; 3.
 DR SMART: SM00261; FU; 1.
 DR SMART: SM00219; TYRK; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00239; RECEPTOR_TYR_KIN_II; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR Transferrase: Tyrosine-protein kinase: Receptor; Transmembrane;
 KM Glycoprotein: ATP-binding; Phosphorylation; Signal.
 FT SIGNAL 1 29
 FT CHAIN 30 716
 FT PROPEP 717 720
 FT CHAIN 721 1363
 FT DOMAIN 721 928
 FT TRANSMEM 929 949
 FT DOMAIN 950 1363
 FT DOMAIN 994 1283
 FT NP_BIND 1000 1008
 FT BINDING 1028 1028
 FT ACT_SITE 1148 1148
 FT MOD_RES 1174 1174
 FT CARBOHYD 51 51
 FT CARBOHYD 97 97
 FT CARBOHYD 137 137
 FT CARBOHYD 278 278
 FT CARBOHYD 483 483
 FT CARBOHYD 599 599
 FT CARBOHYD 617 617
 FT CARBOHYD 665 665
 FT CARBOHYD 666 666
 FT CARBOHYD 711 711
 FT CARBOHYD 732 732
 FT CARBOHYD 736 736
 FT CARBOHYD 743 743
 FT CARBOHYD 816 816
 FT CARBOHYD 885 885
 FT CARBOHYD 898 898
 SO SEQUENCE 1363 AA; 154104 MW; 238120B4EAB1ED65 CRC64;

Query Match 15.0%; Score 342.5; DB 1; Length 1363;
 Best Local Similarity 28.3%; Pred. No. 8e-19;
 Matches 106; Conservative 40; Mismatches 124; Indels 105; Gaps 20;

QY 9 WGLL-----LALLPCGASVQVCTGDMKRLPASPTHLDMRLHYQCGVQVQNIETL 63
 DB 12 MAALTLVIGLGLVPSNGEYICDSMDIRNR-----VSNLRQL-ENCTVIEGYLQI- 61
 QY 64 YLPTNASLSFLDIOEVQGVLLAHNOVROVP-----LQRLR-----IYAGTQ 106
 DB 62 -----LLIDPAEBODSGIAPFNLVETIDFFLYRVRGLTNSELFPNLAVIRGTN 112
 QY 107 LFDNTALVALDNGDPLNNTPTVGTASPGGLRELSLEILKGVLIQNPOLCYQDPT 166
 DB 113 LF-FNALVLFVEMLD-----MOKIGLYSLQNIITRGVRIEKNKNTLCYLDPT 156
 QY 167 ILWKDIF--HKNNQALATLIDITNRSRAC--HPCSPMK-----GSRCMGESSEDCOSLT 216

DB 157 IDMSFIAESGYSNN-----FIVDNREEECVAFCCGRCRIKHPLVLDICMAE--EHCQKVC 210
 QY 217 RTVAGGCARCKRGLPLPDCDCECAACTGPKHSDCLACLFHFNISGICELCHPVLVYNT 276
 DB 211 PESCLGNCR---DGISGCCHEHCIGCGDPTERDVCACRYFVHNSGCLLCQCPDYQYK 266
 QY 277 D---TFESMPNPEGRY--TFGASCVCAPVNYSTVGS---CTLVCPLHNGEVTAE 327
 DB 267 DRRCITFEECPTNTNSWVKLHHRKCIPECFSGY--TTDINNRLCT----- 310
 QY 328 TQRCF-KCSKPCARG 341
 DB 311 --EEGCGCKPSCKG 323

RESULT 15
 INSR_DROME
 ID INSR_DROME STANDARD; PRT; 2146 AA.
 AC P09208; Q24089; Q24023;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Insulin-like receptor precursor (EC 2.7.1.112).
 GN INR OR INR-A OR DIR-A.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95354555; PubMed=7628438;
 RA Fernandez R., Tabarini D., Azpiroz N., Frasch M., Schlessinger J.;
 RT "The Drosophila insulin receptor homolog: a gene essential for
 RT embryonic development encodes two receptor isoforms with different
 RT signaling potential.";
 RL EMBO J. 14:3373-3384(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95181404; PubMed=7876183;
 RA Ruan Y., Chen C., Cao Y., Garofalo R.S.;
 RT "The Drosophila insulin receptor contains a novel carboxyl-terminal
 RT extension likely to play an important role in signal transduction.";
 RL J. Biol. Chem. 270:4236-4243(1995).
 RN [3]
 RP SEQUENCE OF 652-1749 FROM N.A.
 RX STRAIN=Oregon-R; TISSUE=Embryo;
 MEDLINE=87100165; PubMed=3099787;
 RA Nishida Y., Hata M., Nishizuka Y., Rutter W.J., Ebina Y.;
 RT "Cloning of a Drosophila cDNA encoding a polypeptide similar to the
 RT human insulin receptor precursor.";
 RL Biochem. Biophys. Res. Commun. 141:474-481(1986).
 RN [4]
 RP SEQUENCE OF 1297-1595 FROM N.A.
 RX MEDLINE=86259667; PubMed=3014506;
 RA Petruzzelli L., Herrera R., Arenas-Garcia R., Fernandez R.,
 RA Birnbaum M.J., Rosen O.M.;
 RT "Isolation of a Drosophila genomic sequence homologous to the kinase
 RT domain of the human insulin receptor and detection of the
 RT phosphorylated Drosophila receptor with an anti-peptide antibody.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:4710-4714(1986).
 CC -1- FUNCTION: THIS RECEPTOR PROBABLY BINDS AN INSULIN RELATED PROTEIN
 CC AND HAS A TYROSINE-PROTEIN KINASE ACTIVITY (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBUNIT: TETRAMER OF 2 ALPHA AND 2 BETA CHAINS LINKED BY DISULFIDE
 CC BONDS. THE ALPHA CHAINS CONTRIBUTE TO THE FORMATION OF THE LIGAND-
 CC BINDING DOMAIN, WHILE THE BETA CHAINS CARRY THE KINASE DOMAIN.
 CC -1- SUBCELLULAR LOCATION: TYPE I membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN
 CC RECEPTOR SUBFAMILY.

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:31:26 ; Search time 46.2751 Seconds
(without alignments)
1865.663 Million cell updates/sec

Title: US-09-234-208b-2
Perfect score: 2287
Sequence: 1 METALCRWGLLALLPPGA.....VGRGPPDAHVAVLNRYEG 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_21: *
1: sp._archaea: *
2: sp._bacteria: *
3: sp._fungi: *
4: sp._human: *
5: sp._invertebrate: *
6: sp._mammal: *
7: sp._mhc: *
8: sp._organelle: *
9: sp._phage: *
10: sp._plant: *
11: sp._rodent: *
12: sp._virus: *
13: sp._vertebrate: *
14: sp._unclassified: *
15: sp._virus: *
16: sp._bacteria: *
17: sp._archae: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2275	99.5	419	4 Q9UK79	Q9UK79 homo sapien
2	1716.5	75.1	1259	6 O18735	O18735 canis fam11
3	794	34.7	527	13 Q90836	Q90836 gallus galli
4	789	34.5	643	11 Q9ERV6	Q9ERV6 mus musculu
5	789	34.5	655	11 Q9WVF5	Q9WVF5 mus musculu
6	789	34.5	1210	11 Q9EP98	Q9EP98 mus musculu
7	781.5	34.2	1209	11 Q9GX70	Q9GX70 ratu8 norv
8	773	33.8	478	11 Q9SE50	Q9SE50 ratu8 norv
9	734	32.1	331	4 Q9BUD7	Q9BUD7 homo sapien
10	723	31.6	149	6 Q9BG66	Q9BG66 oryctola8us
11	693	30.3	1165	13 Q9YH40	Q9YH40 xiphophoru8
12	662.5	29.0	1328	13 P79754	P79754 figu rubrip
13	647	28.3	599	13 Q9PSH2	Q9PSH2 gallu8 galli
14	571.5	25.0	1433	5 Q9B1H9	Q9B1H9 anophele8 g
15	463.5	20.3	1137	13 Q9W6F6	Q9W6F6 gallu8 galli
16	404.5	17.7	150	6 Q9BG64	Q9BG64 oryctola8us

17	402.5	17.6	1368	5 Q23821	Q23821 caenorhabd
18	395	17.3	151	6 Q9BG65	Q9BG65 oryctola8us
19	363.5	15.9	366	5 Q26569	Q26569 schistosoma
20	363.5	15.9	1717	5 Q26566	Q26566 schistosoma
21	353.5	15.5	1193	5 Q9Y1X8	Q9Y1X8 ephydattia f
22	331	14.5	334	5 Q26567	Q26567 schistosoma
23	331	14.5	342	5 Q26568	Q26568 schistosoma
24	300.5	13.1	1418	13 Q8UW83	Q8UW83 paratichy
25	300	13.1	1472	5 Q9U5A8	Q9U5A8 bombyx mori
26	298.5	13.1	1671	5 Q9N1V5	Q9N1V5 biophalari
27	298	13.0	1358	13 Q73798	Q73798 xenopus lae
28	298	13.0	1418	13 Q93457	Q93457 scophthalmu
29	294	12.9	2144	5 Q9VD94	Q9VD94 drosophila
30	284	12.4	1362	13 Q9PV24	Q9PV24 xenopus lae
31	280	12.2	1749	5 Q8T0W6	Q8T0W6 echinococu
32	278	12.2	1412	13 Q8UW84	Q8UW84 paratichy
33	275.5	12.0	1369	13 Q8UW86	Q8UW86 paratichy
34	273.5	12.0	89	11 Q88459	Q88459 mus musculu
35	272	11.9	469	11 Q63721	Q63721 ratu8 norv
36	270	11.8	410	11 Q63720	Q63720 ratu8 norv
37	256.5	11.2	1368	13 Q8UW85	Q8UW85 paratichy
38	252.5	11.0	1371	11 Q9QVW4	Q9QVW4 ratu8 sp.
39	232	10.1	1245	13 Q9YGH8	Q9YGH8 scophthalmu
40	230	10.1	946	5 Q9VJ04	Q9VJ04 drosophila
41	210	9.2	868	5 Q9VEE2	Q9VEE2 drosophila
42	203	8.9	1843	5 Q968Y9	Q968Y9 caenorhabd
43	203	8.9	1845	5 Q16131	Q16131 caenorhabd
44	191	8.4	82	6 Q9NOK4	Q9NOK4 sus scrofa
45	189	8.3	131	5 Q9B1H6	Q9B1H6 anophele8 g

ALIGNMENTS

RESULT 1
Q9UK79 PRELIMINARY; PRT; 419 AA.

AC Q9UK79;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Herstatin.
GN HER-2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99415951; PubMed=10485918;
RA Doherty J.K., Bond C., Jardim A., Adelman J.P., Clinton G.M.;
RT "The HER-2/neu receptor tyrosine kinase gene encodes a secreted
RT autoinhibitor.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:10869-10874(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Doherty J.K., Clinton G.M., Adelman J.P., Evans A.J., Henner W.D.;
RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF177761; AAD56009.2; -
DR InterPro: IPR000494; EGFR_L-domain.
DR InterPro: IPR002174; Furin-like.
DR Pfam: PF00757; Furin-like: 1.
DR Pfam: PF01030; Recep_L-domain: 1.
DR SMART; SM00261; FU; 1.
SQ SEQUENCE 419 AA: 45472 MW: FECLBE347E2D030C CRC64;

Query Match 99.5%; Score 2275; DB 4; Length 419;
Best Local Similarity 99.5%; Pred. No. 3e-200; Indels 0; Gaps 0;

Matches 417; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 METALCRWGLLALLPPGAAS7QVCTGTDMLRLPASPETHLMDLRHLVYGGQVVGNTL 60
DB 1 METALCRWGLLALLPPGAAS7QVCTGTDMLRLPASPETHLMDLRHLVYGGQVVGNTL 60

Oy	61	ELTYPTNASTLSFLDIOIOVOGYVLIANOVROVLOEIRVSTOTOFEDNYATAVDNG	120
Db	61	ELTYPTNASTLSFLDIOIOVOGYVLIANOVROVLOEIRVSTOTOLFEDNYATAVDNG	120
Oy	121	DLPLNTTFTVTGASPEGLELOLRSLTEILKGGVLIQORNPOLCYDITLIMKDIFFKNNOLA	180
Db	121	DLPLNTTFTVTGASPEGLELOLRSLTEILKGGVLIQORNPOLCYDITLIMKDIFFKNNOLA	180
Oy	181	LTLIDITNRSRACHPCSPBCKMSRCMGESSEPCOSITLTVCGAGCARCKGPIPTCCHEOC	240
Db	181	LTLIDITNRSRACHPCSPBCKMSRCMGESSEPCOSITLTVCGAGCARCKGPIPTCCHEOC	240
Oy	241	AAAGCTGPRXSDCLACLFHNHSGICEHLCPALVTYNTDITFESMPNEBGRYTFGASCVTACP	300
Db	241	AAAGCTGPRXSDCLACLFHNHSGICEHLCPALVTYNTDITFESMPNEBGRYTFGASCVTACP	300
Oy	301	YNYLSTDVGSGCTIVCLPINOEVTAEDGTQOREKSKSPCARSTHSLLRPAVAVPLMOP	360
Db	301	YNYLSTDVGSGCTIVCLPINOEVTAEDGTQOREKSKSPCARSTHSLLRPAVAVPLMOP	360
Oy	361	GRAHVYLSFLRPSMDLVSAFYSPLAPLSPISVSPISVSGRGDPPAHAAVAVNLSREG	419
Db	361	GRAHVYLSFLRPSMDLVSAFYSPLAPLSPISVSPISVSGRGDPPAHAAVAVNLSREG	419

RESULT 2			
ID	018735	PRELIMINARY:	PRT: 1259 AA.
AC	018735		
DT	01-JAN-1998 (TREMBLrel. 05, Created)		
DT	01-JAN-1998 (TREMBLrel. 05, Last sequence update)		
DE	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
OS	Erbb-2		
OC	Canis familiaris (Dog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Yokota H.		
RT	"CDNA cloning of erbb-2 from canine mammary gland.";		
RL	Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AB008451; BAA23127.1; -		
DR	HSSP; P11362; IFCG.		
DR	InterPro: IPR002048; EF-hand.		
DR	InterPro: IPR000494; EGFR_L_domain.		
DR	InterPro: IPR000719; Euk_pkinase.		
DR	InterPro: IPR002174; Furin-kinase.		
DR	InterPro: IPR001245; Tyr-kinase.		
DR	InterPro: IPR004019; YLP_motif.		
DR	Pfam; PF00757; Furin-like; 1.		
DR	Pfam; PF00069; pkinase; 1.		
DR	Pfam; PF01030; Recep_L_domain; 2.		
DR	Pfam; PF02757; YLP. 2.		
DR	ProDom; PD000001; Euk_pkinase; 1.		
DR	SMART; SM00261; FU; 3.		
DR	SMART; SM00219; Tyrc; 1.		
DR	PROSITE; PS00018; EF_HAND; UNKNOWN_1.		
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.		
DR	PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.		
DR	PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.		
KW	ATP-binding; Transferase; Tyrosine-protein kinase.		
QW	SEQUENCE 1259 AA; 137989 MW; E37364D9C4ACD46 CRC64;		

Query Match 75.1%; Score 1716.5; DB 6; Length 1259;
 Best Local Similarity 82.2%; Pred. No. 2e-148;
 Matches 324; Conservative 14; Mismatches 49; Indels 7; Gaps 2.

Oy	61	ELTYLPTNASSTFLODIOEVOGYVLIAHNOYQVPLQRLRIVGTSTGFENYALAVLDNG	120
Dp	61	ELTYLPANASSTFLODIOEVOGYVLIAHNSQYRPLQRLRIVGTSTGFEDNYALAVLDNG	120
Oy	121	DPLNNTPTVTASGGGLRELDLSLTTELKGGVLIQXNPOLCYQDTILMKDIFHKNNOLA	180
Dp	121	DPLEGGIPAPAAAGGGLRELDLSLTTELKGGVLIQSPOLCHQDTILMKDIVEHKNNOILA	180
Oy	181	LTLIDTNRSRACHCQSPCKSRGCMGSSSEDCOSLRTVAGGCAKRGKPLPDDCCHEQC	240
Dp	181	LTLIDTNRFSACPPCSPACKAHCMGSSSDGQSLRTVACGGCAKRGKGPLDDCCHEQC	240
Oy	241	AAAGCTGPRHSCCLAFHFNHSGITELHCPALYNTYNTTFESMPRPERRTYFGASCTYACP	300
Dp	241	AAAGCTGPRHSCCLAFHFNHSGITELHCPALYNTYNTTFESMPRPERRTYFGASCTYSCP	300
Oy	301	YNYLSTDVGSCITLCPHNDQVTAEDGTQREKCSKPCAR-----GTHSLLPAPAAVPYV	355.5
Dp	301	YNYLSTDVGSCITLCPHNDQVTAEDGTQREKCSKPCARCYGGLGMEHLREVRAVTSAN	360
Oy	356	LRMQPG--PAHPVLSFLRPSMDVLSATYSIPLAP	387
Dp	361	IQERAGCKKRTGSLAFEPSEDFGDPASNTAFLQP	394

ID	Q90836	PRELIMINARY:	PT:	527 AA.
AC	Q90836:			
DT	01-NOV-1996 (TReMBLrel. 01, Created)			
DT	01-NOV-1996 (TReMBLrel. 01, Last sequence update)			
DE	01-JUN-2002 (TReMBLrel. 21, Last annotation update)			
DE	EGF/TFG-alpha receptor precursor.			
GN	C-ERBB.			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OX	Gallus.			
NCBI	taxid=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92123214; Pubmed=1733751;			
RA	Flickinger T.W., Malbie N.J., Kung H.-J.;			
RT	"An alternatively processed mRNA from the avian c-erbB gene encodes a			
RT	soluble, truncated form of the receptor that can block ligand-			
RT	dependent transfection."			
RL	MoM. Cell. Biol. 12:883-893(1992).			
EMBL	M77637; AAA84759.1; -			
DR	InterPro: IPR000494; EGF_L_domain.			
DR	InterPro: IPR002174; Furin-like.			
DR	Pfam: PF00757; Furin-like. 1.			
DR	Pfam: PF01030; Recep_Ldomain; 2.			
SMART	SM00261; FU; 2.			
DR	Receptor; Signal.			
FT	SIGNAL	1	28	POTENTIAL.
FT	CHAIN	29	527	EGF/TFG-ALPHA RECEPTOR.
SEQUENCE	527 AA; 58353 MW; 764564ABC095298 CRC64;			

[illegible]

OY 186 -TNR8RACHPGSCPMKSGKSCWGESSSDCOSLRPTVACGCA-RCKGRLPTDCHCCAG 243
 Db 190 FASNSLSSCKKPNCTEDHCKMGAGEONCOTTLKRVICADQCGSGCKRKSVPDCCHNCAAG 249
 OY 244 CTGPKHSDCLACLFHNHSGICELHCPALVTYVTTDFPESMNPBEGKRTFCASCVTACPYNY 303
 Db 250 CTGPRESDCLACRKRFRDADTKCOTCEPVLVNPPTTYQMDVNPBEGKRSFGATCYBCEPHNY 309
 OY 304 ISTDVSGCTLCVPLNHQVEYTAEDGDRCKSGKSPCAR 340
 Db 310 VYTDHGSCVRSCTDYTEV-FENGVRKCKKCDGLCSK 345
 RESULT 4
 OQERV6 PRELIMINARY; PRT; 643 AA.
 AC OQERV6;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Epidermal growth factor receptor isoform 2.
 GN EGFR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H/101, 129/SVJ, AND 129/SVEVYAC;
 RA Relter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,
 RA Sinclair C.S., Parsall R.S., Green P.J., Yee D., Lampard A.L.,
 RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,
 RA Mahle N.J.;
 RT "Comparative genomic sequence analysis and isolation of human and
 RT mouse alternative Egfr transcripts encoding truncated receptor
 RT isoforms";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF275366; AAC28046.1; JOINED.
 DR EMBL: AF275365; AAC28046.1; JOINED.
 DR MGD: MGI:95294; Egfr.
 DR InterPro: IPR000345; Cyto_c_heme_bind.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR002174; Furin-like.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR SMART: SM00261; FU; 4.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 KW Receptor.
 SO SEQUENCE 643 AA; 71476 MW; DEF22002c84911B1 CRC64;
 Query Match 34.5%; Score 789; DB 11; Length 643;
 Best Local Similarity 46.3%; Pred. No. 9.8e-64;
 Matches 156; Conservative 43; Mismatches 116; Indels 22; Gaps
 OY 11 LLLALTPGAA-STVOYCTGMDKRLRLASPTTHLDMLRKLKYGOGVNGNLELTPTN 68
 Db 14 LITLALCAAGGALAEKKVVOGTSNRKLTOGLTFEDHLSLDORMNCEVVLGNLEITTVORN 73
 OY 69 ASLSFLQDIQEVGYVLLAHNOVROVPLORLIRVGTOLFEEDNYALAVLNDGDLNTPR 128
 Db 74 YDLSEFLTKIQEYVGVLLALNTVERIPLENLQIRGNALYENTYVALALUS----- 124
 OY 129 VTGASPGGRLRELQSLSTLEILKGVLIQRNPCLCYODITLMDI----FHKNNQLALTLI 184
 Db 125 -YGTNRGTGRLREIPMRNLDLLEILGAVRFSSNNPILCNMDITIQMRDIOVQNFMSNMSDL--- 180
 OY 185 DTNR8RACHPGSCPMKSGKSCWGESSSDCOSLRPTVACGCA-RCKGRLPTDCHCCAG 243
 Db 181 -QSHHSSGCPKCDPSCPNSGSCWGGEGENCKLKITIACQCGSHRCKGRSRSDCCHNCAAG 239
 OY 244 CTGPKHSDCLACLFHNHSGICELHCPALVTYVTTDFPESMNPBEGKRTFCASCVTACPYNY 303

[illegible]

RA Petch L.A., Harris J., Raymond V.W., Blasband A.J., Lee D.C.,
 RA Earp H.S.;
 RT "A truncated, secreted form of the epidermal growth factor receptor is
 RT encoded by an alternatively spliced transcript in normal rat tissue";
 RL Mol. Cell. Biol. 10:2973-2982(1990).
 RN [12]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FISHER; TISSUE-LIVER;
 RA Petch L.A.;
 RL Submitted (NOV-1991) to the EMBL/GenBank/DBJ databases.
 RN [13]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FISHER; TISSUE-LIVER;
 RA Gutteridge K., Dawson T.L., Earp H.S.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: M37394; AF14008.1; -
 DR HSSP: P11362; 1FGK.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00261; FU; 3.
 DR SMART: SM00219; TYRK; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR ATP-binding; Receptor; Transferase; Tyrosine-protein kinase.
 SQ SEQUENCE 1209 AA; 134891 MW; 96FEETFE6CC1B7773 CRC64;

Query Match 34.2%; Score 781.5; DB 11; Length 1209;
 Best Local Similarity 46.0%; Pred. No. 1,1e-62;
 Matches 157; Conservative 44; Mismatches 117; Indels 23; Gaps 7;

OY 3 LALACRMGLLALLPAGA--ASTOYCTGTDMKRLRPASPEHLDMLRHLXYGCGQYVGNLE 61
 15 LALACGAG-----GALEKKVCCOSTNRNLQGLTFEDHFLSLQRMNCEVGLNLE 66
 OY 62 LTYLPTNASLFDLQIEVGVYLIANOVROYPLQRLRYRGQLFEDNYALAVLDNGD 121
 DB 67 ITVYQRYVDSFLKTIQEVAGVYLIANTVERIPLENLQIIRGALXENYALAVLSN-- 124
 OY 122 PLNNTPYVTASFGLELDELBSLTELKGVLIQRRNPOLCYODTILMKDIFIKNNOLAL 181
 DB 125 -----YGTNKTGLRELPMRNLOELIGAVRFSNNPILCNMETTIQWRDIV-QDVFLSN 175
 OY 182 TLIDTNRN-RACHPCSPMKSGRCMGSSSEDCSLPFTVCAGGA-RCKGPLPTDCHEQ 239
 DB 176 MSMDVORHLTGCRKDPSCNGSCMGEGENCQKLTIKIAQOCSTRKGRSPSDCCHNQ 235
 OY 240 CAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMNPREGRTFGASCYTAC 259
 DB 236 CAAGCTGPRSDCLVCHRFDEATCKDTCRPLMLYNFTYQMDVNPBGKYSFGATCVKCC 255
 OY 300 PNYVLTSDVGSCTLVCPRLHNOEVTAEEDGTORCEKCSKPCAR 340
 DB 296 PRNYVYTDHGSVCVACGPDYEV-EEDGVSKCKKCDGPCRK 335

RESULT 8
 Q9ESFO PRELIMINARY; PRT; 478 AA.
 AC Q9ESFO;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-OCT-2001 (Tremblrel. 18, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Epidermal growth factor receptor related protein.
 GN ERK.
 OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=101116;
 RN [11]
 RP SEQUENCE FROM N.A.
 RC TISSUE-GASTRO-DUODENAL MUCOUS;
 RA Yu Y., Moshier J.A., Majumdar A.P.N.;
 RT "Cloning of a novel EGFR-related peptide: A putative negative
 RT regulator of EGFR";
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF187818; AAG17037.2; -
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR002174; Furin-like.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR Receptor.
 KW SEQUENCE 478 AA; 53233 MW; CF873A8376C519E5 CRC64;
 SQ SEQUENCE 331 AA; 36489 MW; 45B8BEE683FE7E8 CRC64;

Query Match 33.8%; Score 773; DB 11; Length 478;
 Best Local Similarity 45.7%; Pred. No. 2e-62;
 Matches 154; Conservative 43; Mismatches 118; Indels 22; Gaps 6;

OY 11 LLLALLPAGA--STOYCTGTDMKRLRPASPEHLDMLRHLXYGCGQYVGNLELYLPTN 68
 DB 14 LTLALCAAGALBEKKVCCOSTNRNLQGLTFEDHFLSLQRMNCEVGLNLEITYVQRN 73
 OY 69 ASLSFLQDIEVGVYLIANOVROYPLQRLRYRGQLFEDNYALAVLDNGDPLNNTTP 128
 DB 74 YDSFLKTIQEVAGVYLIANTVERIPSEDLQIIRGALXENYALAVLSN----- 124
 OY 129 VTGASPGLELDELBSLTELKGVLIQRRNPOLCYODTILMKDI---FKNNOLATLI 184
 DB 125 -YGTNKTGLRELPMRNLOELIGAVRFSNNPILCNMETTIQWRDIVQNVFMSMSMDL--- 180
 OY 185 DYNRSRACHPCSPMKSGRCMGSSSEDCSLPFTVCAGGA-RCKGPLPTDCHEQCAAG 243
 DB 181 -QSHPSKPCRCDPSCNGSCMGEGENCQKLTIKIAQOCSTRKGRSPSDCCHNQCAAG 239
 OY 244 CTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMNPREGRTFGASCYTACYN 303
 DB 240 CTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMNPREGRTFGASCYTACYN 299
 OY 304 LSTPDVGSCTLVCPRLHNOEVTAEEDGTORCEKCSKPCAR 340
 DB 300 VTDHGSVCVACGPDYEV-EEDGIRKCKKCDGPCRK 335

RESULT 9
 Q9BUD7 PRELIMINARY; PRT; 331 AA.
 AC Q9BUD7;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Similar to v-erb-B2 avian erythroblastic leukemia viral oncogene
 DE homolog 3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
 OX NCBI_TaxID=9606;
 RN [11]
 RP SEQUENCE FROM N.A.
 RC TISSUE-PLACENTA;
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC002706; AAH02706.1; -
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR002174; Furin-like.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF01030; Recep_L_domain; 1.
 DR SMART: SM00261; FU; 2.
 SQ SEQUENCE 331 AA; 36489 MW; 45B8BEE683FE7E8 CRC64;

```

Query Match          32.1%; Score 734; DB 4; Length 331;
Best Local Similarity 44.0%; Pred. No. 4,7e-59;
Matches 147; Conservative 49; Mismatches 120; Indels 18; Gaps 8;

QY 10 GLLALLPPGAA---SPVGTGTMKRLPASPETHLDMLRHLVGGVOCVGNLELTLP 67
DB 11 GLLFSLARSGEVSNAVCPTGLNGLSDAENQYOTLKIKERCYVWGNLEIVLTGH 70
QY 68 NASISFLQDIOEVGYVLIANOVROYPLQRLIRVGTQLEFEDNYALAVLNDGPNLNT 127
DB 71 NADLSFLQWIREVGYVLIANNEFSTLPRLNLRVVGSTQYDGFALFVW----LNYNT 125
QY 128 PYVGASGGLRELQRLSTLILKGVLIQNRNPOLCYDTILMKDIFKNNQALTLIDTN 187
DB 126 ----NSSHALRQLRLQTLTILSGVYIEKNDKCHMDTIDMRDVRDRD---AEIVYKD 178
QY 188 RSRACHPCSPMGKSGRCMGSESEDCOSLRTVCAGGC-ARCKGFLPTDCHEOCACGCTG 246
DB 179 NGRSCPCPCHEVCKG-KRMVGSGSEDCQTLRTKTCAPQCNCHGCFGNPNQCHDECAGCSCG 237
QY 247 PKHSDDLACHFNHSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYACPYNTLST 306
DB 238 PODTDFACPHFNDSGACVRCPOPLVYKRLTFQLEPNPHTKYQYGGVCAVACPHNFV-V 296
QY 307 DVGSCTLVCPHNGEYVAEDGTQCEKSKPCAR 340
DB 297 DQTSVCARACPPDKMEVD-KNGLKMCPCPGGLCPK 329

RESULT 10
Q9BG66 PRELIMINARY; PRT; 149 AA.
AC Q9BG66;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)
DE Receptor tyrosine kinase ErbB2 (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Klonisch T., Wolf P., Hombach-Klonisch S., Vogt S., Kuechenhoff A.,
RA "ErbB genes and epidermal growth factor- (EGF-) like ligands in the
RA peri-implantation rabbit uterus and blastocyst."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF333178; AAK14371.1; -.
DR InterPro: IPR002174; Furin-like.
DR Pfam: PF00757; Furin-like; 1.
DR SMART; SM00261; FU; 2.
KW Kinase.
FT NON_TER 1
FT NON_TER 149
SQ SEQUENCE 149 AA; 16240 MW; 7CB3792A54FC49BA CRC64;

Query Match          31.6%; Score 723; DB 6; Length 149;
Best Local Similarity 81.2%; Pred. No. 1.8e-58;
Matches 121; Conservative 8; Mismatches 20; Indels 0; Gaps 0;

QY 158 NPOLCYDTILMKDIFKNNQALTLIDTNRSRACHPCSPMGKSGRCMGSESEDCOSLTR 217
DB 1 NPQFCYDTILMGEFSTRITTSWPSSTRINASRARCPSCAPCAOSGCGSPEDCCOSLTR 60
QY 218 TYVAGCCARCKGFLPTDCHEOCACGCTGPKHSDDLACHFNHSGICELHCPALVYNTD 277
DB 61 TTAGGCGARCKGFLPTDCHEOCACGCTGPKHSDDLACHFNHSGICELHCPALVYNTD 120
QY 278 TFESMPNPEGRTYFGASCYACPYNTLST 306
DB 121 TFESMPNPEGRTYFGASCYACPYNTLST 149

```

```

RESULT 11
Q9YH40 PRELIMINARY; PRT; 1165 AA.
AC Q9YH40;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Receptor tyrosine kinase proto-oncogene.
GN XMRK.
OS Xiphophorus xiphidium.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Acanthomorphia; Acanthopterygii; Percormorpha; Atherinomorpha;
OC Cyprinodontiformes; Poeciliidae; Xiphophorus.
OX NCBI_TaxID=8086;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RIO PURIFICATION;
RX MEDLINE=98241172; PubMed=9582016;
RA Dimitrijevic N., Winkler C., Wellbrock C., Gomez A., Duschl J.,
RA Altschmid J., Scharl M.;
RT "Activation of the xmrk proto-oncogene of Xiphophorus by
RT overexpression and mutational alterations."
RL Oncogene 16:1681-1690(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-RIO PURIFICATION;
RA Scharl M.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; U53471; AAD10500.2; -.
DR HSSP; P11362; IFGK.
DR InterPro: IPR000345; CytC_heme_bind.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF00069; Kinase; 1.
DR Pfam: PF01030; Recept_L_domain; 2.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Euk_pkinase; 1.
DR SMART; SM00261; FU; 3.
DR DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 2.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 1165 AA; 129614 MW; 7F7EB38D871A7AE CRC64;

Query Match          30.3%; Score 693; DB 13; Length 1165;
Best Local Similarity 42.8%; Pred. No. 1.3e-54;
Matches 148; Conservative 46; Mismatches 134; Indels 18; Gaps 9;

QY 1 MELALACRMGLLALLPPG-AAST---QVCTGTMKRLPASPETHLDMLRHLVGGCOY 55
DB 4 LELLT-----LILLLSIGRCSTDPDRKVCQGTSMQWFM---LDNHYLKMKMYGCNV 56
QY 56 VQGNLELTLPYNASLFLQDIOEVGYVLIANOVROYPLQRLIRVGTQLEFEDNYALA 115
DB 57 VLENLETTYQENQDLSFLQSDIEVGYVLIANNEFSTLPRLNLRVVGSTQYDGFALFVW 116
QY 116 VLNDGPNLNTTQVTVGASPGGLRELQRLSTLILKGVLIQNRNPOLCYDTILMKDIFK 175
DB 117 VMSNYOK-NESSP--DYVQYGLKQLOLSNLTETILSGVYKSHNPDLNVEETIMWDIVDK 173
QY 176 NNQALTLIDTNRSRACHPCSPMGKSGRCMGSESEDCOSLRTVCAGGC-ARCKGFLPTD 234
DB 174 TSNPTNMLIPHAERQCGKCDPCVNGSCWAPRGHCQKTKLLCAQCNRRRCGRKPID 233

```


Db 534 PGPDHMKCAFIDGPHCVKACPAGV 559

RESULT 14

09BIH9 PRELIMINARY; PRT: 1433 AA.
 AC 09BIH9;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Putative epidermal growth factor receptor (Fragment).
 GN EGFR.
 OS Anopheles gambiae (African malaria mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
 OC Anopheles.
 OX NCBI_TaxId=7165;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SU4;
 RA Lycett G.J.;
 RT "Cloning, expression and localisation of the Anopheles gambiae
 epidermal growth factor receptor";
 RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AJ301653; CAC35008.1; -;
 DR HSP; P11362; 1FGK.
 DR InterPro: IPR000345; Cytochrome_b5.
 DR InterPro: IPR000494; EGFR_Ldomain.
 DR InterPro: IPR000719; Euk_kinase.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF00069; kinase; 1.
 DR Pfam: PF01030; Recep_Ldomain; 2.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Euk_kinase; 1.
 DR SMART; SM00261; FU; 7.
 DR SMART; SM00220; S_TKc; 1.
 DR SMART; SM00219; TyKc; 1.
 DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_4.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 KW Receptor.
 FT NON_TER
 SQ SEQUENCE 1433 AA; 159585 MW; E3DD98967724F07 CRC64;

Query Match 25.0%; Score 571.5; DB 5; Length 1433;
 Best Local Similarity 37.2%; Pred. No. 2.4e-43;
 Matches 120; Conservative 47; Mismatches 125; Indels 31; Gaps 7;

OY 26 CIGTGMKLLPASPETHMLRLHYGCGVOGNLELYPTMASISFLDIOEYGVYL 85
 DB 1 CIGTGMKSVLPANREHYHYNLDRYNTCTYVDGNETITWIONTTDINFTQHREVGVYL 60
 OY 86 IAHNOYQVPLRLRIVRSTOLF-----EDNYALAVLDNGDPLNNTT PYTGASPGGLREL 140
 DB 61 ISLYDLPQVILRLQIIRRTFTFKLNKMEAVGLFV-----SFSIMNTL 104
 OY 141 QLRSLTEILKGVLIQRNPOLCYDITLKKDI-FHKNNOLATLIDTNSRACHPCSPMC 199
 DB 105 ELPLADLIDIGSVGFNNYLCMKSIMNEEILLAPOTSOMOTYFNSSPERVCPCHPSC 164
 OY 200 KSRKMGSESDCCSLTRVCAAGCA--RCKGLPLPDCCHBQCAAGCTGPKHSDCLACH 257
 DB 165 EVG-CWGEAGHNCQRFKSLNCSPOCSQGRCFGKPRECHLFCAGGCTGPTQSDCLACKN 223
 OY 258 FHHSGICEHCPALATVYNDTFEESMPNPGRYTFGASCYTACPYNLTSDVSGCTLVCP 317
 DB 224 FYDDGVCCKQECPPMOIYNTFNFMENPDGKATYAGTCVARKCP-EHLLDNGACVAKCKR 282
 OY 318 HNOEYTAEDGTORCEKSKPCAR 340

Db 283 GRMPONSE-----CVPCKRGVCPK 300

RESULT 15

09W6F6 PRELIMINARY; PRT: 1137 AA.
 AC 09W6F6;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Receptor tyrosine kinase (Fragment).
 GN ERBB4.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxId=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HINDRAIN;
 RX MEDLINE=99263203; PubMed=10328884;
 RA Dixon M., Lumsden A.;
 RT "Distribution of neuregulin-1 (nrg1) and erbB4 transcripts in
 embryonic chick hindbrain";
 RL Mol. Cell. Neurosci. 13:237-258(1999).
 DR EMBL; AF121963; AAD31764.1; -;
 DR HSP; P11362; 1FGK.
 DR InterPro: IPR000494; EGFR_Ldomain.
 DR InterPro: IPR000719; Euk_kinase.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR001368; TNFR_c6.
 DR InterPro: IPR001245; Tyr_kinase.
 DR InterPro: IPR004019; YLP_motif.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF00069; kinase; 1.
 DR Pfam: PF01030; Recep_Ldomain; 1.
 DR Pfam: PF02757; YLP; 2.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Euk_kinase; 1.
 DR SMART; SM00261; FU; 3.
 DR SMART; SM00219; TyKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
 KW kinase; Tyrosine-protein kinase.
 FT NON_TER
 SQ SEQUENCE 1137 AA; 127927 MW; 4D616436F87DC84F CRC64;

Query Match 20.3%; Score 463.5; DB 13; Length 1137;
 Best Local Similarity 44.8%; Pred. No. 1.4e-33;
 Matches 86; Conservative 24; Mismatches 73; Indels 9; Gaps 5;

OY 161 LCYQDTILKKDIFHKNNOLATLIDTNSRACHPCSPMKGSRMGSESDCCSLTRIVC 220
 DB 3 LCFADTIHQDIIVRNFWASNFTLVPTNGSSCGGRCHKSCGTG-RCMGPTNHCQTILTKVC 61
 OY 221 AGGC-ARCKGLPLPDCCHBQCAAGCTGPKHSDCLACHHNSGICELHCPALVYNTDF 279
 DB 62 ABOCDGRCTGYPVSDCHRECAAGCGSPKDTDCFACMNFNDGACVTOCPQFFVNPPTF 121
 OY 280 ESMNPNGRYTFGASCVTCPYNYLSTDVSGCTLVCPLNHOEYTAEDGTORCEKSKPCA 339
 DB 122 QLEHNNNAKTYTGATCVKCPHNFV-VDSSECVRACPSKMEV-EEHGKMKRPTDIDCP 179
 OY 340 R-----GTHSL 346
 DB 180 KACDGIQTSGLV 191

Search completed: March 4, 2003, 12:36:02
 Job time: 50.2751 secs

1

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:32:36 ; Search time 22.7169 seconds
(without alignments)
542.689 million cell updates/sec

Title: US-09-234-208B-2
Perfect score: 2287
Sequence: 1 METALACRMGLLALLPPGA.....VGRGPDPAHVAVNLRYEG 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfillseq1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2287	100.0	419	4	US-09-630-155-2	Sequence 2, Appl1
2	1878	82.1	782	2	US-09-146-283-4	Sequence 4, Appl1
3	1878	82.1	782	2	US-08-579-823A-4	Sequence 4, Appl1
4	1878	82.1	782	4	US-09-344-195-4	Sequence 4, Appl1
5	1878	82.1	1255	1	US-08-467-083-68	Sequence 68, Appl1
6	1878	82.1	1255	1	US-08-414-417B-68	Sequence 68, Appl1
7	1878	82.1	1255	2	US-08-484-438-8	Sequence 68, Appl1
8	1878	82.1	1255	2	US-08-486-348A-68	Sequence 68, Appl1
9	1878	82.1	1255	2	US-08-625-101-2	Sequence 68, Appl1
10	1878	82.1	1255	2	US-08-468-545B-68	Sequence 68, Appl1
11	1878	82.1	1255	3	US-08-356-786-2	Sequence 2, Appl1
12	1878	82.1	1255	3	US-08-466-680B-68	Sequence 68, Appl1
13	1769	77.4	624	3	US-08-422-108-1	Sequence 1, Appl1
14	1769	77.4	624	4	US-08-422-734-1	Sequence 1, Appl1
15	793	34.7	644	1	US-08-336-708A-9	Sequence 9, Appl1
16	793	34.7	1210	2	US-08-484-438-7	Sequence 7, Appl1
17	793	34.7	1210	2	US-08-475-035-4	Sequence 4, Appl1
18	775	33.9	911	2	US-08-484-438-10	Sequence 10, Appl1
19	775	33.9	1058	2	US-08-484-438-4	Sequence 4, Appl1
20	775	33.9	1308	2	US-08-484-438-2	Sequence 2, Appl1
21	773	33.8	478	4	US-09-570-895-4	Sequence 2, Appl1
22	735.5	32.2	1342	1	US-07-978-895-4	Sequence 9, Appl1
23	735.5	32.2	1342	2	US-08-484-438-9	Sequence 4, Appl1
24	735.5	32.2	1342	2	US-08-473-119-4	Sequence 4, Appl1
25	735.5	32.2	1342	2	US-08-475-352-4	Sequence 4, Appl1
26	734	32.1	1343	6	5183884-4	Patent No. 5183884
27	493	21.6	97	1	US-08-421-356-3	Sequence 3, Appl1

28	493	21.6	97	4	US-09-046-783-3	Sequence 3, Appl1
29	418	18.3	79	4	US-09-630-155-1	Sequence 1, Appl1
30	264.5	11.6	1382	2	US-08-737-715-2	Sequence 2, Appl1
31	264.5	11.6	1382	4	US-09-457-040B-7	Sequence 7, Appl1
32	257.5	11.3	316	3	US-08-746-559A-4	Sequence 4, Appl1
33	257.5	11.3	1367	2	US-08-249-687C-2	Sequence 2, Appl1
34	257.5	11.3	1367	2	US-08-625-819-2	Sequence 2, Appl1
35	257.5	11.3	1367	4	US-08-746-559A-2	Sequence 2, Appl1
36	257.5	11.3	1367	4	US-08-864-641B-18	Sequence 18, Appl1
37	241.5	10.6	486	3	US-08-746-559A-5	Sequence 5, Appl1
38	210.5	9.2	383	4	US-08-857-076-105	Sequence 105, App
39	203	8.9	1724	4	US-08-857-076-12	Sequence 12, Appl
40	196	8.6	366	4	US-08-857-076-103	Sequence 103, App
41	184.5	8.1	370	4	US-08-857-076-104	Sequence 104, App
42	147.5	6.4	381	4	US-08-857-076-106	Sequence 106, App
43	142	6.2	1940	2	US-08-644-271-30	Sequence 30, Appl
44	142	6.2	1940	2	US-09-077-955-34	Sequence 34, Appl
45	131.5	5.7	799	2	US-08-525-940-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-09-630-155-2
; Sequence 2, Application US/09630155
; Patent No. 6414130
GENERAL INFORMATION:
; APPLICANT: Doherty, Joni Kristin and Gail M. Clinton
; TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESSES:
; ADDRESSER: DAVIS WRIGHT TREMAINE LLP
; STREET: 1501 Fourth Avenue, 2600 Century Square
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98101
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: PC compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: Word
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/630,155
; FILING DATE: 16-Jan-2001
; CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
; NAME: Davison, Barry L.
; REGISTRATION NUMBER: 47,309
; REFERENCE/DOCKET NUMBER: 49321-10
TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206 628-7621
; TELEFAX: 206 628-7699
INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: polypeptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-630-155-2

Query Match 100.0%; Score 2287; DB 4; Length 419;
Best Local Similarity 100.0%; Pred. No. 2,8e-192;
Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 METALACRMGLLALLPPGAASVQCTGTDMLRLPASPETHIDMLRHLYOGGVQGNL 60
|||||
DB 1 METALACRMGLLALLPPGAASVQCTGTDMLRLPASPETHIDMLRHLYOGGVQGNL 60
|||||
QY 61 ELTYPLTNASTLSPDIOEVOGVVLLAHNNOVROVPIORLRITVNGTOLFEDNVALAVLDNG 120

```
|||||
Db 61 ELTYLPTNASTSLFLODIOEQVGYVLIAHNOVROVPLQRLIRVGTQLEFEDNVALAVLDNG 120
QY 121 DPLNNTPTVYGASPGGLRELOLRSLTEILKGVLIQRNPQLCYQDTILMKDIFHKNNOLA 180
Db 121 DPLNNTPTVYGASPGGLRELOLRSLTEILKGVLIQRNPQLCYQDTILMKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMCKGSRCKWGESSEDDCSLTRTYCAGGCARCKGPLPTDCCHEQC 240
Db 181 LTLIDNRSRACHPCSPMCKGSRCKWGESSEDDCSLTRTYCAGGCARCKGPLPTDCCHEQC 240
QY 241 AACCTGPKHSDCLACLFHFNHSGICELHCPALVYNTDTFESMPNPEGRTYFGASCVTACP 300
Db 241 AACCTGPKHSDCLACLFHFNHSGICELHCPALVYNTDTFESMPNPEGRTYFGASCVTACP 300
QY 301 YNLTSDVGSCTLVCPHLNDEVTAEDGTQRCCKSPKACNGTSLPRPAVYVPLRMOP 360
Db 301 YNLTSDVGSCTLVCPHLNDEVTAEDGTQRCCKSPKACNGTSLPRPAVYVPLRMOP 360
QY 361 GPAHPVLSFLRPSMDLVSAFYSLPLAPLSPTSVPISPVSVGRGPDPAHVAVNLRSYEG 419
Db 361 GPAHPVLSFLRPSMDLVSAFYSLPLAPLSPTSVPISPVSVGRGPDPAHVAVNLRSYEG 419
```

RESULT 2

```
US-09-146-283-4
; Sequence 4, Application US/09146283
```

```
; Patent No. 5976546
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Laus, Reiner
```

```
; APPLICANT: Ruegg, Curtis L.
```

```
; APPLICANT: Wu, Hongyu
```

```
; TITLE OF INVENTION: Immunostimulatory Compositions
```

```
; NUMBER OF SEQUENCES: 10
```

```
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Dehlinger & Associates
```

```
; STREET: 350 Cambridge Ave. Suite 250
```

```
; CITY: Palo Alto
```

```
; STATE: CA
```

```
; COUNTRY: USA
```

```
; ZIP: 94306
```

```
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: floppy disk
```

```
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: Patentin Release #1.0, Version #1.25
```

```
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/09/146,283
```

```
; FILING DATE: 03-SEPT-1998
```

```
; CLASSIFICATION: 536
```

```
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Judge, Linda R.
```

```
; REGISTRATION NUMBER: 42,702
```

```
; REFERENCE/DOCKET NUMBER: 7636-0010.21
```

```
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: 650-324-0880
```

```
; TELEFAX: 650-324-0960
```

```
; INFORMATION FOR SEQ. ID NO: 4:
```

```
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 782 amino acids
```

```
; TYPE: amino acid
```

```
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: protein
```

```
; HYPOTHETICAL: NO
```

```
; ORGANISM: homo sapiens
```

```
; INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
```

```
US-09-146-283-4
```

```
Query Match 82.1%; Score 1878; DB 2; Length 782;  
Best Local Similarity 83.0%; Pred. No. 4e-156;  
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
```

```
QY 1 METALALCRWGLLLALLPAGASTOVCTGTDMLRLPASPETHLDMRLHLYXGCGQVQGNL 60  
Db 1 METALALCRWGLLLALLPAGASTOVCTGTDMLRLPASPETHLDMRLHLYXGCGQVQGNL 60  
QY 61 ELTYLPTNASTSLFLODIOEQVGYVLIAHNOVROVPLQRLIRVGTQLEFEDNVALAVLDNG 120  
Db 61 ELTYLPTNASTSLFLODIOEQVGYVLIAHNOVROVPLQRLIRVGTQLEFEDNVALAVLDNG 120  
QY 121 DPLNNTPTVYGASPGGLRELOLRSLTEILKGVLIQRNPQLCYQDTILMKDIFHKNNOLA 180  
Db 121 DPLNNTPTVYGASPGGLRELOLRSLTEILKGVLIQRNPQLCYQDTILMKDIFHKNNOLA 180  
QY 181 LTLIDNRSRACHPCSPMCKGSRCKWGESSEDDCSLTRTYCAGGCARCKGPLPTDCCHEQC 240  
Db 181 LTLIDNRSRACHPCSPMCKGSRCKWGESSEDDCSLTRTYCAGGCARCKGPLPTDCCHEQC 240  
QY 241 AACCTGPKHSDCLACLFHFNHSGICELHCPALVYNTDTFESMPNPEGRTYFGASCVTACP 300  
Db 241 AACCTGPKHSDCLACLFHFNHSGICELHCPALVYNTDTFESMPNPEGRTYFGASCVTACP 300  
QY 301 YNLTSDVGSCTLVCPHLNDEVTAEDGTQRCCKSPKACNGTSLPRPAVYVPLRMOP 360  
Db 301 YNLTSDVGSCTLVCPHLNDEVTAEDGTQRCCKSPKACNGTSLPRPAVYVPLRMOP 360  
QY 361 GPAHPVLSFLRPSMDLVSAFYSLPLAPLSPTSVPISPVSVGRGPDPAHVAVNLRSYEG 419  
Db 361 GPAHPVLSFLRPSMDLVSAFYSLPLAPLSPTSVPISPVSVGRGPDPAHVAVNLRSYEG 419
```

RESULT 3

```
US-08-579-823A-4  
; Sequence 4, Application US/08579823A
```

```
; Patent No. 6080409
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Laus, Reiner
```

```
; APPLICANT: Ruegg, Curtis L.
```

```
; APPLICANT: Wu, Hongyu
```

```
; TITLE OF INVENTION: Immunostimulatory Composition and Method
```

```
; NUMBER OF SEQUENCES: 10
```

```
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Dehlinger & Associates
```

```
; STREET: 350 Cambridge Ave. Suite 250
```

```
; CITY: Palo Alto
```

```
; STATE: CA
```

```
; COUNTRY: USA
```

```
; ZIP: 94306
```

```
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: floppy disk
```

```
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: Patentin Release #1.0, Version #1.25
```

```
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/579,823A
```

```
; FILING DATE: 03-DEC-1998
```

```
; CLASSIFICATION: 536
```

```
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Judge, Linda R.
```

```
; REGISTRATION NUMBER: 42,702
```

```
; REFERENCE/DOCKET NUMBER: 7636-0010
```

```
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: 650-324-0880
```

```
; TELEFAX: 650-324-0960
```

```
; INFORMATION FOR SEQ. ID NO: 4:
```

```
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 782 amino acids
```

```
; TYPE: amino acid
```

```
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: protein
```

```
; HYPOTHETICAL: NO
```



```

; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
US-08-579-823A-4

Query Match      82.1%; Score 1878; DB 3; Length 782;
Best local Similarity 83.0%; Pred. No. 4e-156;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

OY 1 METALACRMGLLALLPRAASTOVCTGDMKRLRPASPEHIDMLRHLYGCGVVOGML 60
    |||
DB 1 METALACRMGLLALLPRAASTOVCTGDMKRLRPASPEHIDMLRHLYGCGVVOGML 60

OY 61 ELTYLPTNLSLFLDIOIEVOGVYLLAHNOVROVPLQRLRIYNGTOLFEDNYALAVLDNG 120
    |||
DB 61 ELTYLPTNLSLFLDIOIEVOGVYLLAHNOVROVPLQRLRIYNGTOLFEDNYALAVLDNG 120

OY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQRPOLCYODTILMKDIFHKNNOLA 180
    |||
DB 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQRPOLCYODTILMKDIFHKNNOLA 180

OY 181 LTLIDNRSRACHPCSPMKGSRCKWSESSEDCQSLRTVACAGCARKGRLPTDCCHEOC 240
    |||
DB 181 LTLIDNRSRACHPCSPMKGSRCKWSESSEDCQSLRTVACAGCARKGRLPTDCCHEOC 240

OY 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMPNPGRTTFGASCYTACP 300
    |||
DB 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMPNPGRTTFGASCYTACP 300

OY 301 YNLTSDVGSCTLVCPDLHNOEYTAEDGTORCEKSKPCAR-----GTHSLRPRAAVPV 355
    |||
DB 301 YNLTSDVGSCTLVCPDLHNOEYTAEDGTORCEKSKPCAR-----GTHSLRPRAAVPV 355

OY 356 LRMQPG--PAHNVLSFLRPSMDLVSFAFYLPLAPLSPVPI-----SPVSVGRGPD 405
    |||
DB 356 LRMQPG--PAHNVLSFLRPSMDLVSFAFYLPLAPLSPVPI-----SPVSVGRGPD 405

OY 406 --PDAHVAVNLSRYEG 419
    |||
DB 406 --PDAHVAVNLSRYEG 419

OY 418 SLPDLVSFQNLQVIRG 433
    |||
DB 418 SLPDLVSFQNLQVIRG 433

RESULT 4
US-09-344-195-4
; Sequence 4, Application US/09344195
; Patent No. 6210662
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; Ruegg, Curtis L.
; Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/344,195
; FILING DATE: 24-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
```

```

; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 782 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-344-195-4

Query Match      82.1%; Score 1878; DB 4; Length 782;
Best local Similarity 83.0%; Pred. No. 4e-156;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

OY 1 METALACRMGLLALLPRAASTOVCTGDMKRLRPASPEHIDMLRHLYGCGVVOGML 60
    |||
DB 1 METALACRMGLLALLPRAASTOVCTGDMKRLRPASPEHIDMLRHLYGCGVVOGML 60

OY 61 ELTYLPTNLSLFLDIOIEVOGVYLLAHNOVROVPLQRLRIYNGTOLFEDNYALAVLDNG 120
    |||
DB 61 ELTYLPTNLSLFLDIOIEVOGVYLLAHNOVROVPLQRLRIYNGTOLFEDNYALAVLDNG 120

OY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQRPOLCYODTILMKDIFHKNNOLA 180
    |||
DB 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQRPOLCYODTILMKDIFHKNNOLA 180

OY 181 LTLIDNRSRACHPCSPMKGSRCKWSESSEDCQSLRTVACAGCARKGRLPTDCCHEOC 240
    |||
DB 181 LTLIDNRSRACHPCSPMKGSRCKWSESSEDCQSLRTVACAGCARKGRLPTDCCHEOC 240

OY 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMPNPGRTTFGASCYTACP 300
    |||
DB 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMPNPGRTTFGASCYTACP 300

OY 301 YNLTSDVGSCTLVCPDLHNOEYTAEDGTORCEKSKPCAR-----GTHSLRPRAAVPV 355
    |||
DB 301 YNLTSDVGSCTLVCPDLHNOEYTAEDGTORCEKSKPCAR-----GTHSLRPRAAVPV 355

OY 356 LRMQPG--PAHNVLSFLRPSMDLVSFAFYLPLAPLSPVPI-----SPVSVGRGPD 405
    |||
DB 356 LRMQPG--PAHNVLSFLRPSMDLVSFAFYLPLAPLSPVPI-----SPVSVGRGPD 405

OY 406 --PDAHVAVNLSRYEG 419
    |||
DB 406 --PDAHVAVNLSRYEG 419

OY 418 SLPDLVSFQNLQVIRG 433
    |||
DB 418 SLPDLVSFQNLQVIRG 433

RESULT 5
US-08-467-083-68
; Sequence 68, Application US/08467083
; Patent No. 5726023
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; DIS: Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN
; FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/NEU ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,083
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/414,417
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; TELEX: 3723836 SEEDANBERRY
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-08-467-083-68
;
Query Match      82.1%; Score 1878; DB 1: Length 1255;
Best Local Similarity 83.0%; Pred. No. 7.5e-156;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
;
OY 1 MELAALCRMGLLALLPPGAASVQVCTGTDMLRLPASPEHIDMLRHLYOGGQVQGNL 60
DB 1 MELAALCRMGLLALLPPGAASVQVCTGTDMLRLPASPEHIDMLRHLYOGGQVQGNL 60
OY 61 ELTYLPTNLSLFLDIOIEVQGYVLLAHNOVROPRLQRLRIYVGTQLFEDNYALAYLDNG 120
DB 61 ELTYLPTNLSLFLDIOIEVQGYVLLAHNOVROPRLQRLRIYVGTQLFEDNYALAYLDNG 120
OY 121 DPLNNTPTVYGASPGGLREQLRSLEILKGVLIQIRNPOLCYQDTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVYGASPGGLREQLRSLEILKGVLIQIRNPOLCYQDTILMKDIFHKNNOLA 180
OY 181 LTLIDNRSRACHPCSPMCKGSRMCWGESSEDCSLTRTYCAGGCARCKGPLPTDCCHQC 240
DB 181 LTLIDNRSRACHPCSPMCKGSRMCWGESSEDCSLTRTYCAGGCARCKGPLPTDCCHQC 240
OY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMNPBGRTYTFGASCTYAC 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMNPBGRTYTFGASCTYAC 300
OY 301 YNLSTDVGSCTLVCPPLHNOEVAEDGTORCEKSKPCAR-----GTHSLPRPAVP 355
DB 301 YNLSTDVGSCTLVCPPLHNOEVAEDGTORCEKSKPCAR-----GTHSLPRPAVP 355
OY 356 LRMQPG--PAHPVLSFLRPSMDLVSFSLPLAPLSPTSVPI-----SPVSVGRGPD 405
DB 356 LRMQPG--PAHPVLSFLRPSMDLVSFSLPLAPLSPTSVPI-----SPVSVGRGPD 405
OY 406 --PDAHVAVNLSRYEG 419
DB 406 --PDAHVAVNLSRYEG 419
OY 418 SLPDLSVFQNLQVIRG 433
DB 418 SLPDLSVFQNLQVIRG 433
;
RESULT 6
US-08-414-417B-68
; Sequence 66, Application US/08414417B
; Patent No. 5801005
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
```

```

; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/414,417B
; FILING DATE: 31-MAR-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-08-414-417B-68
;
Query Match      82.1%; Score 1878; DB 1: Length 1255;
Best Local Similarity 83.0%; Pred. No. 7.5e-156;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
;
OY 1 MELAALCRMGLLALLPPGAASVQVCTGTDMLRLPASPEHIDMLRHLYOGGQVQGNL 60
DB 1 MELAALCRMGLLALLPPGAASVQVCTGTDMLRLPASPEHIDMLRHLYOGGQVQGNL 60
OY 61 ELTYLPTNLSLFLDIOIEVQGYVLLAHNOVROPRLQRLRIYVGTQLFEDNYALAYLDNG 120
DB 61 ELTYLPTNLSLFLDIOIEVQGYVLLAHNOVROPRLQRLRIYVGTQLFEDNYALAYLDNG 120
OY 121 DPLNNTPTVYGASPGGLREQLRSLEILKGVLIQIRNPOLCYQDTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVYGASPGGLREQLRSLEILKGVLIQIRNPOLCYQDTILMKDIFHKNNOLA 180
OY 181 LTLIDNRSRACHPCSPMCKGSRMCWGESSEDCSLTRTYCAGGCARCKGPLPTDCCHQC 240
DB 181 LTLIDNRSRACHPCSPMCKGSRMCWGESSEDCSLTRTYCAGGCARCKGPLPTDCCHQC 240
OY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMNPBGRTYTFGASCTYAC 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMNPBGRTYTFGASCTYAC 300
OY 301 YNLSTDVGSCTLVCPPLHNOEVAEDGTORCEKSKPCAR-----GTHSLPRPAVP 355
DB 301 YNLSTDVGSCTLVCPPLHNOEVAEDGTORCEKSKPCAR-----GTHSLPRPAVP 355
OY 356 LRMQPG--PAHPVLSFLRPSMDLVSFSLPLAPLSPTSVPI-----SPVSVGRGPD 405
DB 356 LRMQPG--PAHPVLSFLRPSMDLVSFSLPLAPLSPTSVPI-----SPVSVGRGPD 405
OY 406 --PDAHVAVNLSRYEG 419
DB 406 --PDAHVAVNLSRYEG 419
OY 418 SLPDLSVFQNLQVIRG 433
DB 418 SLPDLSVFQNLQVIRG 433
;
RESULT 7
US-08-484-438-8
; Sequence 8, Application US/08484438
; Patent No. 5811098
```

Patent No. 5811098 5780031
GENERAL INFORMATION:
APPLICANT: Plowman, Gregory D.
APPLICANT: Culouscou, Jean-Michel
APPLICANT: Shoyab, Mohammed
APPLICANT: Slegail, Clay B.
APPLICANT: Hellstr m, Ingegerd
APPLICANT: Hellstr m, Karl E.
TITLE OF INVENTION: HERA HUMAN RECEPTOR TYROSINE KINASE
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,438
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/323,442
FILING DATE: 14-OCT-1994
APPLICATION NUMBER: US 08/150,704
FILING DATE: 10-NOV-1993
CLASSIFICATION: 530
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/981,165
FILING DATE: 24-NOV-1992
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5624-230
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-484-438-8

Query Match 82.1%; Score 1878; DB 2; Length 1355;
Best Local Similarity 83.0%; Pred. No. 7.5e-156;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 MELAAICRMGLALLALPPGASTOVCTGTDMLRLPASPTHTDMLRHLYOGCOVVOGNL 60
DB 1 MELAAICRMGLALLALPPGASTOVCTGTDMLRLPASPTHTDMLRHLYOGCOVVOGNL 60
QY 61 ELTYLPTNASTSLFDIOEVOGVVLIAHNOVROVPLQRLRIYVGTOLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASTSLFDIOEVOGVVLIAHNOVROVPLQRLRIYVGTOLFEDNYALAVLDNG 120
QY 121 DPLNNTPTVVGASPGGLRELQRLSLTEILKGVLIQRNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVVGASPGGLRELQRLSLTEILKGVLIQRNPOLCYODTILMKDIFHKNNOLA 180
QY 121 DPLNNTPTVVGASPGGLRELQRLSLTEILKGVLIQRNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVVGASPGGLRELQRLSLTEILKGVLIQRNPOLCYODTILMKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMCKGSRCKWESSSDCOSLRTYVAGGACARCKGLPLDCCHEOC 240
DB 181 LTLIDNRSRACHPCSPMCKGSRCKWESSSDCOSLRTYVAGGACARCKGLPLDCCHEOC 240

QY 241 AAGCTGPKHSKDLACLAHFHNHSGICELCHPALVYNTDFFESMPNDEGRYTFGASCYTACP 300
DB 241 AAGCTGPKHSKDLACLAHFHNHSGICELCHPALVYNTDFFESMPNDEGRYTFGASCYTACP 300
QY 301 YNYLSTDVSGCTLYCPLHNOEYTAEDGTQRCCKSPCAR----GTHSLPRAAVVP 355
DB 301 YNYLSTDVSGCTLYCPLHNOEYTAEDGTQRCCKSPCARVCYGLGMEHLREVAAYTSAN 360
QY 356 LRMPG--PAHPVLSFLRPMDLVSAFSLPLAPLSTSPV-----SPSVGKRPD 405
DB 361 IQEFAGCKKIFGSLAFIPESFDGPASNT--APLOEQLQVETLEIRGYLIISAMPD 417
QY 406 --PDAAHVAVNLRYEG 419
DB 418 SLPLSLVFNQLQVIRG 433

RESULT 8
US-08-486-348A-68
Sequence 68, Application US/08486348A
Patent No. 584538
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Disis, Mary L.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,348A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-486-348A-68

Query Match 82.1%; Score 1878; DB 2; Length 1355;
Best Local Similarity 83.0%; Pred. No. 7.5e-156;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 MELAAICRMGLALLALPPGASTOVCTGTDMLRLPASPTHTDMLRHLYOGCOVVOGNL 60
DB 1 MELAAICRMGLALLALPPGASTOVCTGTDMLRLPASPTHTDMLRHLYOGCOVVOGNL 60
QY 61 ELTYLPTNASTSLFDIOEVOGVVLIAHNOVROVPLQRLRIYVGTOLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASTSLFDIOEVOGVVLIAHNOVROVPLQRLRIYVGTOLFEDNYALAVLDNG 120
QY 121 DPLNNTPTVVGASPGGLRELQRLSLTEILKGVLIQRNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVVGASPGGLRELQRLSLTEILKGVLIQRNPOLCYODTILMKDIFHKNNOLA 180


```

OY 1 MELAALCRWGILLALPPGAASOVCTGDMKRLPASPETHLMDRLHYOGCOVVOGNI 60
    |||
Db 1 MELAALCRWGILLALPPGAASOVCTGDMKRLPASPETHLMDRLHYOGCOVVOGNI 60
OY 61 ELTYLPTNASTSLFDIOIEVOGYVLLAHNQVROPVLPQRLRIYRGTLFEDNYALAVLDNG 120
    |||
Db 61 ELTYLPTNASTSLFDIOIEVOGYVLLAHNQVROPVLPQRLRIYRGTLFEDNYALAVLDNG 120
OY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQRPOLCYODTILMKDIFHKNNOLA 180
    |||
Db 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQRPOLCYODTILMKDIFHKNNOLA 180
OY 181 LFLIDNRRACHPCSPMKSGSRMGESSEDCOSLRTVCAGGACARCKPLPTDCCHEC 240
    |||
Db 181 LFLIDNRRACHPCSPMKSGSRMGESSEDCOSLRTVCAGGACARCKPLPTDCCHEC 240
OY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFTESMPNPEGRTYFGASCYTACP 300
    |||
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFTESMPNPEGRTYFGASCYTACP 300
OY 301 YNLTSDVSGCTLVCPPLHNOEYTAEDGTORCEKSKPCAR-----GTHSLPRPAVPVP 355
    |||
Db 301 YNLTSDVSGCTLVCPPLHNOEYTAEDGTORCEKSKPCAR-----GTHSLPRPAVPVP 355
OY 356 LRMQPG--PAHPVLSFLRPSMDLVSAFYSLPLAPLSPTSVPI-----SPYVGRGPD 405
    |||
Db 356 LRMQPG--PAHPVLSFLRPSMDLVSAFYSLPLAPLSPTSVPI-----SPYVGRGPD 405
OY 406 --PDAHVAVNLSRYEG 419
    |||
Db 406 --PDAHVAVNLSRYEG 419
    |||
Db 418 SLPLDSVFQNLQVIRG 433

```

```

RESULT 11
US-08-356-786-2
; Sequence 2, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppertmann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; TITLE OF INVENTION: Marker
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,786
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/831,967
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: CRP-053
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 2:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-356-786-2

```

Query Match 82.1%; Score 1878; DB 2; Length 1255;

Best Local Similarity 83.0%; Pred. No. 7,5e-156; Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

```

OY 1 MELAALCRWGILLALPPGAASOVCTGDMKRLPASPETHLMDRLHYOGCOVVOGNI 60
    |||
Db 1 MELAALCRWGILLALPPGAASOVCTGDMKRLPASPETHLMDRLHYOGCOVVOGNI 60
OY 61 ELTYLPTNASTSLFDIOIEVOGYVLLAHNQVROPVLPQRLRIYRGTLFEDNYALAVLDNG 120
    |||
Db 61 ELTYLPTNASTSLFDIOIEVOGYVLLAHNQVROPVLPQRLRIYRGTLFEDNYALAVLDNG 120
OY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQRPOLCYODTILMKDIFHKNNOLA 180
    |||
Db 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQRPOLCYODTILMKDIFHKNNOLA 180
OY 181 LFLIDNRRACHPCSPMKSGSRMGESSEDCOSLRTVCAGGACARCKPLPTDCCHEC 240
    |||
Db 181 LFLIDNRRACHPCSPMKSGSRMGESSEDCOSLRTVCAGGACARCKPLPTDCCHEC 240
OY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFTESMPNPEGRTYFGASCYTACP 300
    |||
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFTESMPNPEGRTYFGASCYTACP 300
OY 301 YNLTSDVSGCTLVCPPLHNOEYTAEDGTORCEKSKPCAR-----GTHSLPRPAVPVP 355
    |||
Db 301 YNLTSDVSGCTLVCPPLHNOEYTAEDGTORCEKSKPCAR-----GTHSLPRPAVPVP 355
OY 356 LRMQPG--PAHPVLSFLRPSMDLVSAFYSLPLAPLSPTSVPI-----SPYVGRGPD 405
    |||
Db 356 LRMQPG--PAHPVLSFLRPSMDLVSAFYSLPLAPLSPTSVPI-----SPYVGRGPD 405
OY 406 --PDAHVAVNLSRYEG 419
    |||
Db 406 --PDAHVAVNLSRYEG 419
    |||
Db 418 SLPLDSVFQNLQVIRG 433

```

```

RESULT 12
US-08-466-680B-68
; Sequence 68, Application US/08466680B
; Patent No. 6075122
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,680B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.

```

```

;
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-466-680B-68

Query Match      82.1%; Score 1878; DB 3; Length 1255;
Best Local Similarity 83.0%; Pred. No. 7.5e-156;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 MELAALCRWGLLALLPGGAATQVCTGDMKRLRPAISPETHLDMRLHLYOGCCVQVGNL 60
DB 1 MELAALCRWGLLALLPGGAATQVCTGDMKRLRPAISPETHLDMRLHLYOGCCVQVGNL 60
QY 61 ELTYLPTNASTSLFLQDIOEVGYVLLAHNOVROVPLQRLRIVRGTOLEFEDNVALAVLDNG 120
DB 61 ELTYLPTNASTSLFLQDIOEVGYVLLAHNOVROVPLQRLRIVRGTOLEFEDNVALAVLDNG 120
QY 121 DPLNNTPTVYGASPGGLRELQRLSLEFTEILKGVLIQRPOLCYQDTILMKDIFHKNNQLA 180
DB 121 DPLNNTPTVYGASPGGLRELQRLSLEFTEILKGVLIQRPOLCYQDTILMKDIFHKNNQLA 180
QY 181 LTLIDNRSRACHPCSPMCKSGRCWSESSSDCSLTRTVAGGCAKCKGLPTDCCHEQC 240
DB 181 LTLIDNRSRACHPCSPMCKSGRCWSESSSDCSLTRTVAGGCAKCKGLPTDCCHEQC 240
QY 241 AAGCTPKHSDCLACLFHNSGICELHCPALVTNTDFESMPNBEGRYTFGASCYTACP 300
DB 241 AAGCTPKHSDCLACLFHNSGICELHCPALVTNTDFESMPNBEGRYTFGASCYTACP 300
QY 301 YNTLSTDVGSCTLVCPHLNQEVTAEDEGTOCEKSKPCAR----GTHSLRPPAAVPPV 355
DB 301 YNTLSTDVGSCTLVCPHLNQEVTAEDEGTOCEKSKPCAR----GTHSLRPPAAVPPV 355
QY 356 LRMQPG--PAHPVLSFLRPSMDLVSAFYSLPLAPLSPTSVPI-----SPVSVGSGPD 405
DB 356 LRMQPG--PAHPVLSFLRPSMDLVSAFYSLPLAPLSPTSVPI-----SPVSVGSGPD 405
QY 406 --PDAAVAVNLSTRYEG 419
DB 418 SLRDLVSFQNLQYIRG 433

RESULT 13
US-08-422-108-1
; Sequence 1, Application US/08422108
; Patent No. 6015567
; GENERAL INFORMATION:
; APPLICANT: Hudziak, Robert M.
; APPLICANT: Shepard, H. Michael
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,108
```

```

;
; FILING DATE: 14-Apr-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/355460
; FILING DATE: 13-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/048346
; FILING DATE: 15-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/354319
; FILING DATE: 19-MAY-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 554C2D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 624 amino acids
; TYPE: Amino Acid
; TOPOLOGY: linear
US-08-422-108-1

Query Match      77.4%; Score 1769; DB 3; Length 624;
Best Local Similarity 82.2%; Pred. No. 1e-146;
Matches 341; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 22 STQVCTGTDMLRLPASPETHLDMRLHLYOGCCVQVGNLELYLPTNASTSLFLQDIOEVQ 81
DB 1 STQVCTGTDMLRLPASPETHLDMRLHLYOGCCVQVGNLELYLPTNASTSLFLQDIOEVQ 81
QY 82 GYVLLAHNOVROVPLQRLRIVRGTOLEFEDNVALAVLDNGDPLNNTPTVYGASPGGLRELQ 141
DB 61 GYVLLAHNOVROVPLQRLRIVRGTOLEFEDNVALAVLDNGDPLNNTPTVYGASPGGLRELQ 120
QY 142 LRLTEILKGVLIQRPOLCYQDTILMKDIFHKNNQLATLIDTNRSRACHPCSPMCKG 201
DB 121 LRLTEILKGVLIQRPOLCYQDTILMKDIFHKNNQLATLIDTNRSRACHPCSPMCKG 180
QY 202 SRCWGESSEDCQSLTRTVAGGCAKCKGLPTDCCHEQCAAGCTGKHSCLACLFHNS 261
DB 181 SRCWGESSEDCQSLTRTVAGGCAKCKGLPTDCCHEQCAAGCTGKHSCLACLFHNS 240
QY 262 GICELHCPALVTNTDFESMPNBEGRYTFGASCYTACPNTLSTDVGSCTLVCPHLNQE 321
DB 241 GICELHCPALVTNTDFESMPNBEGRYTFGASCYTACPNTLSTDVGSCTLVCPHLNQE 300
QY 322 VTAEDEGTOCEKSKPCAR----GTHSLRPPAAVPPVPLRMQPG--PAHPVLSFLRPSW 374
DB 301 VTAEDEGTOCEKSKPCARCYGLGMEHLREVAAYVSANQIEFGACKKIFGSLAFPEST 360
QY 375 DLVSIFYSLPLAPLSPTSVPI-----SPVSVGSGPD--PDAAVAVNLSTRYEG 419
DB 361 DGDPAASNT--APLQPEQLOVFETLEITEITYLISAMPDLSPLVSFQNLQYIRG 412

RESULT 14
US-08-422-734-1
; Sequence 1, Application US/08422734
; Patent No. 6333169
; GENERAL INFORMATION:
; APPLICANT: Hudziak, Robert M.
; APPLICANT: Shepard, H. Michael
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
```

```
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPacIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,734
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/422108
FILING DATE: 14-Apr-1995
APPLICATION NUMBER: 08/35460
FILING DATE: 13-Dec-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/048346
FILING DATE: 15-Apr-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/354319
FILING DATE: 19-MAY-1989
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 554C201
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 624 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-422-734-1
```

```
Query Match 77.4% Score 1769; DB 4; Length 624;
Best Local Similarity 82.2%; Pred. No. 1e-146;
Matches 341; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 22 STGVCTGTDMLRLPASPETHLMDLRLHYOGCQVYVGNLELTYLPTNASLFLDIDIEVQ 81
DB 1 STGVCTGTDMLRLPASPETHLMDLRLHYOGCQVYVGNLELTYLPTNASLFLDIDIEVQ 60

QY 82 GYVLIANOVROVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNTPVYTGASPGELRELQ 141
DB 61 GYVLIANOVROVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNTPVYTGASPGELRELQ 120

QY 142 LRSITELKGGVLIORNPOLCYODTILMKDIFHKNNOLATLTDNRSRACHPCSPCKG 201
DB 121 LRSITELKGGVLIORNPOLCYODTILMKDIFHKNNOLATLTDNRSRACHPCSPCKG 180

QY 202 SRMGESSEDCOSLTRVYAGGACARCKPLPTDCHDOCAAGCTGPRHSDCLACIHNHNS 261
DB 181 SRMGESSEDCOSLTRVYAGGACARCKPLPTDCHDOCAAGCTGPRHSDCLACIHNHNS 240

QY 262 GICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACPYNYLSTDVGSCTLYCPRLHNOE 321
DB 241 GICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACPYNYLSTDVGSCTLYCPRLHNOE 300

QY 322 VTAEEDGTORCEKSKPCAR-----GTHSLRPRAAVVPLRMQPG--PAHPVLSFLRPSW 374
DB 301 VTAEEDGTORCEKSKPCARCYGIGMEHLREVRVTSANIOERAGCKKIFGSLAEFLPESF 360

QY 375 DLVSATYSLAPLAPLSPTSPVPI-----SPVSVGRGPD--PDAAHVAVNLSRYEG 419
DB 361 DGDPASNT---APLQPEQLQVFTLEETLEYIVISAMPDLSPLSVQNLQVIRG 412
```

RESULT 15
US-08-336-708A-9

```
Sequence 9, Application US/08336708A
Patent No. 5521295
GENERAL INFORMATION:
APPLICANT: Pacifici, Robert E.
APPLICANT: Thomson, Arlen R.
APPLICANT: Chang, Ming-Shi
TITLE OF INVENTION: Hybrid Receptor Molecules
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 Dehavenland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/336,708A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oleski, Nancy
REFERENCE/DOCKET NUMBER: A-241A
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 644 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-336-708A-9
```

```
Query Match 34.7% Score 793; DB 1; Length 644;
Best Local Similarity 45.3%; Pred. No. 3e-61;
Matches 151; Conservative 50; Mismatches 118; Indels 14; Gaps 4;

QY 11 LLLALPPGAA--STGVCTGTDMLRLPASPETHLMDLRLHYOGCQVYVGNLELTYLPTN 68
DB 14 LLAALCPASRALEEKKYCGTSNKLTLQGTGFEDHFLSLQRFNCEVYVGLSEITTYVGRN 73

QY 69 ASLSFLDIDIEVGYVLIANOVROVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNTP 128
DB 74 YDLSFLKTIDIEVAGVYVLIANOVROVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNTP 126

QY 129 VTGASPGELRELQRLSITELKGGVLIORNPOLCYODTILMKDIFHKNNOLATLTDNTR 188
DB 127 ---ANKTGLKELPRLNIOELIHGAVRSNNPALCNVESIQWRDIVSDPLSNMSDFQNH 183

QY 189 SRACHPCSPCKGSRMGESSEDCOSLTRVYAGGACARCKPLPTDCHDOCAAGCTGP 247
DB 184 LGSQCKCDPSCPNMGSCGAGBENCQKLTITICAQCGRCGRKSPSCCHNCGAAGCTGP 243

QY 248 KHSQCLACILFHNHSGITELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACPYNYLSTD 307
DB 244 RESCDIVCRKFRDEATKDKCPPLMLYNPTTYQMDVBPBGKYSFGATCVKCPRYNYVTD 303

QY 308 VGSCTLYCPRLHNOEVTAEEDGTORCEKSKPCAR 340
DB 304 HGSCVTRACGADSYEM-EDDGVKCKKKEGFCRK 335
```

Search completed: March 4, 2003, 12:37:12
Job time : 26.7169 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:36:06 ; Search time 17.6687 Seconds
(without alignments)
1000.035 Million cell updates/sec

Title: US-09-234-208b-2
Perfect score: 2287
Sequence: 1 METALCRWGLLALLPPGA.....VGRGPRDAHVAVLSRYEG 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188354 seqs, 42170167 residues

Total number of hits satisfying chosen parameters: 188354

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications-AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCT05_PUBCOMB pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep:*
8: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1878	82.1	653	10	US-09-921-161-1
2	1878	82.1	653	9	US-09-854-356-3
3	1878	82.1	712	9	US-09-854-356-7
4	1878	82.1	919	9	US-09-854-356-6
5	1878	82.1	1255	9	US-09-769-508-2
6	1878	82.1	1255	9	US-09-854-356-1
7	1878	82.1	1255	9	US-09-930-125-2
8	1878	82.1	1255	9	US-09-441-411-6
9	1878	82.1	1255	10	US-09-811-113-9
10	1878	82.1	1255	10	US-09-811-113-3
11	1612.5	70.5	479	10	US-09-821-883-5
12	1610.5	70.4	564	10	US-09-821-883-3
13	1610.5	70.4	654	9	US-09-821-883-4
14	1608.5	70.3	654	9	US-09-854-356-8
15	1608.5	70.3	1256	9	US-09-854-356-2
16	1608.5	70.3	1260	9	US-09-870-759-118
17	1607	70.3	555	10	US-09-821-883-1
18	1607	70.3	690	10	US-09-821-883-2
19	1597.5	69.9	1256	9	US-09-854-356-14

20	1587	69.4	289	10	US-09-821-883-23	Sequence 23, Appl
21	987	43.2	191	9	US-09-441-411-9	Sequence 9, Appl
22	793	34.7	1210	10	US-09-725-433-2	Sequence 2, Appl
23	775	33.9	1308	10	US-09-940-101-2	Sequence 2, Appl
24	773	33.8	478	10	US-09-867-521-2	Sequence 4, Appl
25	771	33.7	615	10	US-09-940-101-4	Sequence 4, Appl
26	257.5	11.3	1367	9	US-09-870-759-120	Sequence 120, App
27	210.5	9.2	383	10	US-09-205-658-105	Sequence 105, App
28	210.5	9.2	383	10	US-09-844-353A-105	Sequence 105, App
29	203	8.9	1724	10	US-09-205-658-12	Sequence 12, Appl
30	203	8.9	1724	10	US-09-844-353A-12	Sequence 12, Appl
31	196	8.6	366	10	US-09-205-658-103	Sequence 103, App
32	196	8.6	366	10	US-09-844-353A-103	Sequence 103, App
33	184.5	8.1	370	10	US-09-205-658-104	Sequence 104, App
34	184.5	8.1	370	10	US-09-844-353A-104	Sequence 104, App
35	147.5	6.4	381	10	US-09-205-658-106	Sequence 106, App
36	147.5	6.4	381	10	US-09-844-353A-106	Sequence 106, App
37	142	6.2	1940	9	US-10-016-283-34	Sequence 34, Appl
38	133.5	5.8	420	9	US-09-796-753-86	Sequence 86, Appl
39	130.5	5.7	833	9	US-10-226-296-5	Sequence 5, Appl
40	129.5	5.7	830	9	US-09-870-759-140	Sequence 140, App
41	128.5	5.6	420	9	US-09-905-291A-109	Sequence 109, App
42	128.5	5.6	420	9	US-09-902-853-109	Sequence 109, App
43	128.5	5.6	420	9	US-09-907-824-109	Sequence 109, App
44	128.5	5.6	420	9	US-09-907-841-109	Sequence 109, App
45	128.5	5.6	420	9	US-09-904-011-109	Sequence 109, App

ALIGNMENTS

RESULT 1
US-09-921-161-1
Sequence 1, Application US/09921161
Patent No. US200209062A1
GENERAL INFORMATION:
APPLICANT: Ralph, Peter
TITLE OF INVENTION: ANALYTICAL METHOD
FILE REFERENCE: GENENT. 066A
CURRENT APPLICATION NUMBER: US/09/921,161
CURRENT FILING DATE: 2001-08-01
PRIOR APPLICATION NUMBER: 60/225,433
PRIOR FILING DATE: 2000-08-15
NUMBER OF SEQ ID NOS: 1
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 645
TYPE: PRT
ORGANISM: Homo sapiens
US-09-921-161-1

QY	1	METALCRWGLLALLPPGAASOVCTGTMKRLRASPETHIDMLRHLYOGGVOVGNL	60	82.1%	Score 1878;	DB 10;	Length 645;
					Best local Similarity	83.0%;	Pred. No. 1,6e-130;
		Matches 362;	Conservative	9;	Mismatches	45;	Indels 20;
							Gaps 5;
DB	1	METALCRWGLLALLPPGAASOVCTGTMKRLRASPETHIDMLRHLYOGGVOVGNL	60				
QY	61	ELTYLPTNASTLFLDIOEVGVYLLAHNOVROVPLQRLRTVGTOLFEDNYLAALVDNG	120				
DB	61	ELTYLPTNASTLFLDIOEVGVYLLAHNOVROVPLQRLRTVGTOLFEDNYLAALVDNG	120				
QY	121	DPLANTPTVGTASPGIGRELQRLSLTEILKGVLIQRPOLCYQDTITLWKDITRHKNNQLA	180				
DB	121	DPLANTPTVGTASPGIGRELQRLSLTEILKGVLIQRPOLCYQDTITLWKDITRHKNNQLA	180				
QY	181	LTLIDNRSRACHPCSPMCKSGRCWSESSDQSLRTVAVGACARCKGFLPDDCCEOC	240				
DB	181	LTLIDNRSRACHPCSPMCKSGRCWSESSDQSLRTVAVGACARCKGFLPDDCCEOC	240				
QY	241	AACTGPKHSIDCLAFHFNHSGICELHCPALVYNTDFESMPNPGRYTFGASCVTACP	300				

```
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYTACP 300
QY 301 YNLTSTDVSGCTLVCPPLHNEVTAEDGTQCEKSKPCAR-----GTHSLPRPAAVPVP 355
Db 301 YNLTSTDVSGCTLVCPPLHNEVTAEDGTQCEKSKPCARCYGLGMEHLREVAAYTSAN 360
QY 356 LRMQPG--PAHPVLSTLRPSMDLVSAFYSLPLAPLSPTSVPI-----SPVSVGRGPD 405
Db 361 IQEFAAGCKKIFGSLAFLEPSEFDGDPASNT---APLQPEQLQVETLEITIGYLXISAMPD 417
QY 406 --PDAAVAVNLSRREG 419
Db 418 SLPDLVSFQNLQYIRG 433
```

```
RESULT 2
US-09-854-356-3
; Sequence 3, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
US-09-854-356-3
```

```
Query Match 82.1%; Score 1878; DB 9; Length 653;
Best Local Similarity 83.0%; Pred. No. 1.6e-130;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 MELAALCRWGLLALLPPGAASSTVCTGTDMKRLRPASPEHDMRLHLYOGGCOVQGNL 60
Db 1 MELAALCRWGLLALLPPGAASSTVCTGTDMKRLRPASPEHDMRLHLYOGGCOVQGNL 60
QY 61 ELTYLPTNASTLSPLODIOEVQGYVLIHNOVROVPLQRLRIVRGTOLEFEDNYALAVLDNG 120
Db 61 ELTYLPTNASTLSPLODIOEVQGYVLIHNOVROVPLQRLRIVRGTOLEFEDNYALAVLDNG 120
QY 121 DPLNNTTPVVGASPGGLREQLRSLEILKGVLIQRNPOLCYODTILMKDIFHKNNOLA 180
Db 121 DPLNNTTPVVGASPGGLREQLRSLEILKGVLIQRNPOLCYODTILMKDIFHKNNOLA 180
QY 181 LTLLDTRSRACHPCSPMGKSGRCWGESSEDCOSLRTTVAGGCARCKGPLPTDCHEQC 240
Db 181 LTLLDTRSRACHPCSPMGKSGRCWGESSEDCOSLRTTVAGGCARCKGPLPTDCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYTACP 300
QY 301 YNLTSTDVSGCTLVCPPLHNEVTAEDGTQCEKSKPCAR-----GTHSLPRPAAVPVP 355
Db 301 YNLTSTDVSGCTLVCPPLHNEVTAEDGTQCEKSKPCARCYGLGMEHLREVAAYTSAN 360
QY 356 LRMQPG--PAHPVLSTLRPSMDLVSAFYSLPLAPLSPTSVPI-----SPVSVGRGPD 405
Db 361 IQEFAAGCKKIFGSLAFLEPSEFDGDPASNT---APLQPEQLQVETLEITIGYLXISAMPD 417
```

```
QY 406 --PDAAVAVNLSRREG 419
Db 418 SLPDLVSFQNLQYIRG 433
```

```
RESULT 3
US-09-854-356-7
; Sequence 7, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 712
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: fusion protein
US-09-854-356-7
```

```
Query Match 82.1%; Score 1878; DB 9; Length 712;
Best Local Similarity 83.0%; Pred. No. 1.8e-130;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 MELAALCRWGLLALLPPGAASSTVCTGTDMKRLRPASPEHDMRLHLYOGGCOVQGNL 60
Db 1 MELAALCRWGLLALLPPGAASSTVCTGTDMKRLRPASPEHDMRLHLYOGGCOVQGNL 60
QY 61 ELTYLPTNASTLSPLODIOEVQGYVLIHNOVROVPLQRLRIVRGTOLEFEDNYALAVLDNG 120
Db 61 ELTYLPTNASTLSPLODIOEVQGYVLIHNOVROVPLQRLRIVRGTOLEFEDNYALAVLDNG 120
QY 121 DPLNNTTPVVGASPGGLREQLRSLEILKGVLIQRNPOLCYODTILMKDIFHKNNOLA 180
Db 121 DPLNNTTPVVGASPGGLREQLRSLEILKGVLIQRNPOLCYODTILMKDIFHKNNOLA 180
QY 181 LTLLDTRSRACHPCSPMGKSGRCWGESSEDCOSLRTTVAGGCARCKGPLPTDCHEQC 240
Db 181 LTLLDTRSRACHPCSPMGKSGRCWGESSEDCOSLRTTVAGGCARCKGPLPTDCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYTACP 300
QY 301 YNLTSTDVSGCTLVCPPLHNEVTAEDGTQCEKSKPCAR-----GTHSLPRPAAVPVP 355
Db 301 YNLTSTDVSGCTLVCPPLHNEVTAEDGTQCEKSKPCARCYGLGMEHLREVAAYTSAN 360
QY 356 LRMQPG--PAHPVLSTLRPSMDLVSAFYSLPLAPLSPTSVPI-----SPVSVGRGPD 405
Db 361 IQEFAAGCKKIFGSLAFLEPSEFDGDPASNT---APLQPEQLQVETLEITIGYLXISAMPD 417
QY 406 --PDAAVAVNLSRREG 419
Db 418 SLPDLVSFQNLQYIRG 433
```

```
RESULT 4
US-09-854-356-6
```

```
; Sequence 6, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Cheyssen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 919
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: fusion protein
; OTHER INFORMATION: of ECD and PD of human HER-2/neu
US-09-854-356-6
```

```
Query Match      82.1%; Score 1878; DB 9; Length 919;
Best local Similarity 83.0%; Pred. No. 2,4e-130;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
```

```
QY 1 MELAALCRWGILLALLPFGAASVQVCTGTDMLRLPASPETHIDMLRHLXGCGVQVGNL 60
DB 1 MELAALCRWGILLALLPFGAASVQVCTGTDMLRLPASPETHIDMLRHLXGCGVQVGNL 60
QY 61 ELTYLPTNASLSFLQDIOEVQGYVLLAHNOVQVPLQRLRIYRGTOLEFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIOEVQGYVLLAHNOVQVPLQRLRIYRGTOLEFEDNYALAVLDNG 120
QY 121 DPLNNTPTVGTGASPGGLREQLRLSTLEILKGVLIQIRNPOLCYQDPTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVGTGASPGGLREQLRLSTLEILKGVLIQIRNPOLCYQDPTILMKDIFHKNNOLA 180
QY 121 DPLNNTPTVGTGASPGGLREQLRLSTLEILKGVLIQIRNPOLCYQDPTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVGTGASPGGLREQLRLSTLEILKGVLIQIRNPOLCYQDPTILMKDIFHKNNOLA 180
QY 181 LFLIDNRSRACHPCSPMKGSRGWESSEDCSLRTVACAGCARGKPLPTDCHEQC 240
DB 181 LFLIDNRSRACHPCSPMKGSRGWESSEDCSLRTVACAGCARGKPLPTDCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPGRTYFGASCYTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPGRTYFGASCYTACP 300
QY 301 YNLSLDVSGCTVCPDLHNOEVAEDGTORCEKSKPCAR-----GTHSLRPPRAVPP 355
DB 301 YNLSLDVSGCTVCPDLHNOEVAEDGTORCEKSKPCAR-----GTHSLRPPRAVPP 355
QY 301 YNLSLDVSGCTVCPDLHNOEVAEDGTORCEKSKPCAR-----GTHSLRPPRAVPP 355
DB 301 YNLSLDVSGCTVCPDLHNOEVAEDGTORCEKSKPCAR-----GTHSLRPPRAVPP 355
QY 356 LRMQPG--PAHPVLSFLRPSMDLVSFAFSLPLAPLSPTSVPI-----SPVSGRGPD 405
DB 356 LRMQPG--PAHPVLSFLRPSMDLVSFAFSLPLAPLSPTSVPI-----SPVSGRGPD 405
QY 361 IOEFAGCKKIFGSLAPLPSFDGDPASNT---APLQPEQLQVEETLEITGYLYISAMPD 417
DB 361 IOEFAGCKKIFGSLAPLPSFDGDPASNT---APLQPEQLQVEETLEITGYLYISAMPD 417
QY 406 --PDAAVAVNLSRYEG 419
DB 406 --PDAAVAVNLSRYEG 419
DB 418 SLPDLSVFQNLQVIRG 433
```

```
RESULT 5
US-09-769-508-2
; Sequence 2, Application US/09769508
; Patent No. US2002015527A1
; GENERAL INFORMATION:
; APPLICANT: STUART, SUSAN G.
; APPLICANT: MONAHAN, JOHN J.
; APPLICANT: LANGTON, BEATRICE CLAUDIA
; APPLICANT: HANCOCK, MIRIAM E.C.
; APPLICANT: CHAO, LORRINE A.
```

```
; APPLICANT: BLUFORD, PETER
; TITLE OF INVENTION: C-ERBB-2 EXTERNAL DOMAIN: GP75
; FILE REFERENCE: BEBIO-111-C1
; CURRENT APPLICATION NUMBER: US/09/769,508
; CURRENT FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-769-508-2
```

```
Query Match      82.1%; Score 1878; DB 9; Length 1255;
Best local Similarity 83.0%; Pred. No. 3,5e-130;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
```

```
QY 1 MELAALCRWGILLALLPFGAASVQVCTGTDMLRLPASPETHIDMLRHLXGCGVQVGNL 60
DB 1 MELAALCRWGILLALLPFGAASVQVCTGTDMLRLPASPETHIDMLRHLXGCGVQVGNL 60
QY 61 ELTYLPTNASLSFLQDIOEVQGYVLLAHNOVQVPLQRLRIYRGTOLEFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIOEVQGYVLLAHNOVQVPLQRLRIYRGTOLEFEDNYALAVLDNG 120
QY 121 DPLNNTPTVGTGASPGGLREQLRLSTLEILKGVLIQIRNPOLCYQDPTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVGTGASPGGLREQLRLSTLEILKGVLIQIRNPOLCYQDPTILMKDIFHKNNOLA 180
QY 121 DPLNNTPTVGTGASPGGLREQLRLSTLEILKGVLIQIRNPOLCYQDPTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVGTGASPGGLREQLRLSTLEILKGVLIQIRNPOLCYQDPTILMKDIFHKNNOLA 180
QY 181 LFLIDNRSRACHPCSPMKGSRGWESSEDCSLRTVACAGCARGKPLPTDCHEQC 240
DB 181 LFLIDNRSRACHPCSPMKGSRGWESSEDCSLRTVACAGCARGKPLPTDCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPGRTYFGASCYTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPGRTYFGASCYTACP 300
QY 301 YNLSLDVSGCTVCPDLHNOEVAEDGTORCEKSKPCAR-----GTHSLRPPRAVPP 355
DB 301 YNLSLDVSGCTVCPDLHNOEVAEDGTORCEKSKPCAR-----GTHSLRPPRAVPP 355
QY 301 YNLSLDVSGCTVCPDLHNOEVAEDGTORCEKSKPCAR-----GTHSLRPPRAVPP 355
DB 301 YNLSLDVSGCTVCPDLHNOEVAEDGTORCEKSKPCAR-----GTHSLRPPRAVPP 355
QY 356 LRMQPG--PAHPVLSFLRPSMDLVSFAFSLPLAPLSPTSVPI-----SPVSGRGPD 405
DB 356 LRMQPG--PAHPVLSFLRPSMDLVSFAFSLPLAPLSPTSVPI-----SPVSGRGPD 405
QY 361 IOEFAGCKKIFGSLAPLPSFDGDPASNT---APLQPEQLQVEETLEITGYLYISAMPD 417
DB 361 IOEFAGCKKIFGSLAPLPSFDGDPASNT---APLQPEQLQVEETLEITGYLYISAMPD 417
QY 406 --PDAAVAVNLSRYEG 419
DB 406 --PDAAVAVNLSRYEG 419
DB 418 SLPDLSVFQNLQVIRG 433
```

```
RESULT 6
US-09-854-356-1
; Sequence 1, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Cheyssen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
FEATURE:
OTHER INFORMATION: human HER-2/neu protein
NAME/KEY: DOMAIN
LOCATION: (1)..(653)
OTHER INFORMATION: extracellular domain (ECD)
NAME/KEY: DOMAIN
LOCATION: (676)..(1255)
OTHER INFORMATION: intracellular domain (ICD)
NAME/KEY: DOMAIN
LOCATION: (990)..(1255)
OTHER INFORMATION: phosphorylation domain (PD)
NAME/KEY: DOMAIN
LOCATION: (990)..(1048)
OTHER INFORMATION: fragment of the phosphorylation domain, preferred
OTHER INFORMATION: portion (delta PD)
US-09-854-356-1

Query Match
Best Local Similarity 82.1%; Score 1878; DB 9; Length 1255;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 METALCRWGILLALLPPGAASYVCTGTDMLRLPASETHLDMRLHYQCGQVQGNL 60
DB 1 METALCRWGILLALLPPGAASYVCTGTDMLRLPASETHLDMRLHYQCGQVQGNL 60
QY 61 ELTYLPTNASLSFLDIOEVQGVLLAHNOVQVPLQRLRIYRGQLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLDIOEVQGVLLAHNOVQVPLQRLRIYRGQLFEDNYALAVLDNG 120
QY 121 DPLNNTPTVGTASPGGLRELQRLSLTEILKGVLLIQRNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVGTASPGGLRELQRLSLTEILKGVLLIQRNPOLCYODTILMKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMKKSGRCWGESSEDCQSLRTVCAGGACARCKGPLPTDCCHEQC 240
DB 181 LTLIDNRSRACHPCSPMKKSGRCWGESSEDCQSLRTVCAGGACARCKGPLPTDCCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFESMPNPGRTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFESMPNPGRTFGASCVTACP 300
QY 301 YNLSTDVGSCTLVCPHLNQEVTAEEDGTORCEKSKPCAR-----GTHSLPRPAVAVP 355
DB 301 YNLSTDVGSCTLVCPHLNQEVTAEEDGTORCEKSKPCAR-----GTHSLPRPAVAVP 355
QY 356 LRMQPG--PAHYLSFLRSMDLVSAFYSLPLAPLSPTVPI-----SPVSVGRGPD 405
DB 356 LRMQPG--PAHYLSFLRSMDLVSAFYSLPLAPLSPTVPI-----SPVSVGRGPD 405
QY 406 --PDAHVAVNLRYEG 419
DB 406 --PDAHVAVNLRYEG 419
QY 418 SLPDLVSFONLOYIRG 433
DB 418 SLPDLVSFONLOYIRG 433

RESULT 7
US-09-930-125-2
Sequence 2, Application US/09930125
Publication No. US20020193329A1
GENERAL INFORMATION:
APPLICANT: Hand-Zimmerman, Susan
APPLICANT: Cheever, Martin A.
APPLICANT: Foy, Teresa M.
APPLICANT: Lodes, Michael J.
APPLICANT: Kalos, Michael D.
APPLICANT: McNeill, Patricia D.
APPLICANT: Vedvick, Thomas S.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
FILE REFERENCE: 210121.544
CURRENT APPLICATION NUMBER: US/09/930,125
CURRENT FILING DATE: 2001-08-14
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 3.0
```

```
SEQ ID NO 2
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapien
US-09-930-125-2

Query Match
Best Local Similarity 82.1%; Score 1878; DB 9; Length 1255;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 METALCRWGILLALLPPGAASYVCTGTDMLRLPASETHLDMRLHYQCGQVQGNL 60
DB 1 METALCRWGILLALLPPGAASYVCTGTDMLRLPASETHLDMRLHYQCGQVQGNL 60
QY 61 ELTYLPTNASLSFLDIOEVQGVLLAHNOVQVPLQRLRIYRGQLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLDIOEVQGVLLAHNOVQVPLQRLRIYRGQLFEDNYALAVLDNG 120
QY 121 DPLNNTPTVGTASPGGLRELQRLSLTEILKGVLLIQRNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVGTASPGGLRELQRLSLTEILKGVLLIQRNPOLCYODTILMKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMKKSGRCWGESSEDCQSLRTVCAGGACARCKGPLPTDCCHEQC 240
DB 181 LTLIDNRSRACHPCSPMKKSGRCWGESSEDCQSLRTVCAGGACARCKGPLPTDCCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFESMPNPGRTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFESMPNPGRTFGASCVTACP 300
QY 301 YNLSTDVGSCTLVCPHLNQEVTAEEDGTORCEKSKPCAR-----GTHSLPRPAVAVP 355
DB 301 YNLSTDVGSCTLVCPHLNQEVTAEEDGTORCEKSKPCAR-----GTHSLPRPAVAVP 355
QY 356 LRMQPG--PAHYLSFLRSMDLVSAFYSLPLAPLSPTVPI-----SPVSVGRGPD 405
DB 356 LRMQPG--PAHYLSFLRSMDLVSAFYSLPLAPLSPTVPI-----SPVSVGRGPD 405
QY 406 --PDAHVAVNLRYEG 419
DB 406 --PDAHVAVNLRYEG 419
QY 418 SLPDLVSFONLOYIRG 433
DB 418 SLPDLVSFONLOYIRG 433

RESULT 8
US-09-441-411-6
Sequence 6, Application US/09441411
Publication No. US20030008342A1
GENERAL INFORMATION:
APPLICANT: Scholler, Nathalie B.
APPLICANT: Disis, Mary L.
APPLICANT: Hellstrom, Inggerd
APPLICANT: Hellstrom, Karl Erik
TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
FILE REFERENCE: 730033.409
CURRENT APPLICATION NUMBER: US/09/441,411
CURRENT FILING DATE: 1999-11-16
NUMBER OF SEQ ID NOS: 26
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapiens
US-09-441-411-6

Query Match
Best Local Similarity 82.1%; Score 1878; DB 9; Length 1255;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 METALCRWGILLALLPPGAASYVCTGTDMLRLPASETHLDMRLHYQCGQVQGNL 60
DB 1 METALCRWGILLALLPPGAASYVCTGTDMLRLPASETHLDMRLHYQCGQVQGNL 60
QY 61 ELTYLPTNASLSFLDIOEVQGVLLAHNOVQVPLQRLRIYRGQLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLDIOEVQGVLLAHNOVQVPLQRLRIYRGQLFEDNYALAVLDNG 120
```

```
|||||
Db 61 ELTYLPTNASLSFLQDIQEVQGVYLAHNOVRLPQRLRIYRGTOLEFDNYALAVLDNG 120
|||
Qy 121 DPLNNTPTVYGASPGGLRELOLSLEILKGVLIQIRNPOLCYODITLWKDIFHKNNOLA 180
|||
Db 121 DPLNNTPTVYGASPGGLRELOLSLEILKGVLIQIRNPOLCYODITLWKDIFHKNNOLA 180
|||
Qy 181 LTLIDNRSRACHPCSPMKGSRGMESESDCOSLTRTYCAGGACARCKGPLPTDCCHEQC 240
|||
Db 181 LTLIDNRSRACHPCSPMKGSRGMESESDCOSLTRTYCAGGACARCKGPLPTDCCHEQC 240
|||
Qy 241 AACCTGPKHSDCLACLFHNSGICELHCPALVYNTDFESMPNPGRTYFGASCYTACP 300
|||
Db 241 AACCTGPKHSDCLACLFHNSGICELHCPALVYNTDFESMPNPGRTYFGASCYTACP 300
|||
Qy 301 YNLSLTDVSGCTLVCPPLHNOEYTAEDGTORCEKCSKPCARVCGYGLMEHLREVRATYSAN 360
|||
Db 301 YNLSLTDVSGCTLVCPPLHNOEYTAEDGTORCEKCSKPCARVCGYGLMEHLREVRATYSAN 360
|||
Qy 356 LRMOPG--PAHVPYSLFRPSMDLVSAFYSLPLAPLSPTSVPI-----SPYVGRGPD 405
|||
Db 356 LRMOPG--PAHVPYSLFRPSMDLVSAFYSLPLAPLSPTSVPI-----SPYVGRGPD 405
|||
Qy 406 --PDAAVAVNLSRYEG 419
|||
Db 406 --PDAAVAVNLSRYEG 419
|||
Qy 418 SLDPDLSVFQNLQYIRG 433
|||
Db 418 SLDPDLSVFQNLQYIRG 433
|||
```

RESULT 9
US-09-811-123-9

```
; Sequence 9, Application US/09811123
; Patent No. US20020001587A1
; GENERAL INFORMATION:
; APPLICANT: Sharon Erickson
; APPLICANT: Ralph Schwall
; APPLICANT: Mark Slikowski
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ERBB
; FILE REFERENCE: GENEENT.073A2
; CURRENT APPLICATION NUMBER: US/09/811,123
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/238,327
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 09/602,530
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-123-9
```

Query Match 82.1%; Score 1878; DB 10; Length 1255;
Best Local Similarity 83.0%; Pred. No. 3.5e-130;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

```
Qy 1 MELAALCRMGILLALLPPGAASOVCTGDMKLRLPASPTHLDMLRHL YOGCOVVOGNTL 60
|||
Db 1 MELAALCRMGILLALLPPGAASOVCTGDMKLRLPASPTHLDMLRHL YOGCOVVOGNTL 60
|||
Qy 61 ELTYLPTNASLSFLQDIQEVQGVYLAHNOVRLPQRLRIYRGTOLEFDNYALAVLDNG 120
|||
Db 61 ELTYLPTNASLSFLQDIQEVQGVYLAHNOVRLPQRLRIYRGTOLEFDNYALAVLDNG 120
|||
Qy 121 DPLNNTPTVYGASPGGLRELOLSLEILKGVLIQIRNPOLCYODITLWKDIFHKNNOLA 180
|||
Db 121 DPLNNTPTVYGASPGGLRELOLSLEILKGVLIQIRNPOLCYODITLWKDIFHKNNOLA 180
|||
Qy 181 LTLIDNRSRACHPCSPMKGSRGMESESDCOSLTRTYCAGGACARCKGPLPTDCCHEQC 240
|||
Db 181 LTLIDNRSRACHPCSPMKGSRGMESESDCOSLTRTYCAGGACARCKGPLPTDCCHEQC 240
|||
```

```
Qy 241 AACCTGPKHSDCLACLFHNSGICELHCPALVYNTDFESMPNPGRTYFGASCYTACP 300
|||
Db 241 AACCTGPKHSDCLACLFHNSGICELHCPALVYNTDFESMPNPGRTYFGASCYTACP 300
|||
Qy 301 YNLSLTDVSGCTLVCPPLHNOEYTAEDGTORCEKCSKPCARVCGYGLMEHLREVRATYSAN 360
|||
Db 301 YNLSLTDVSGCTLVCPPLHNOEYTAEDGTORCEKCSKPCARVCGYGLMEHLREVRATYSAN 360
|||
Qy 356 LRMOPG--PAHVPYSLFRPSMDLVSAFYSLPLAPLSPTSVPI-----SPYVGRGPD 405
|||
Db 356 LRMOPG--PAHVPYSLFRPSMDLVSAFYSLPLAPLSPTSVPI-----SPYVGRGPD 405
|||
Qy 406 --PDAAVAVNLSRYEG 419
|||
Db 406 --PDAAVAVNLSRYEG 419
|||
Qy 418 SLDPDLSVFQNLQYIRG 433
|||
Db 418 SLDPDLSVFQNLQYIRG 433
|||
```

RESULT 10
US-09-811-115-3

```
; Sequence 3, Application US/09811115
; Patent No. US20020035736A1
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENEENT.034A
; CURRENT APPLICATION NUMBER: US/09/811,115
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-115-3
```

Query Match 82.1%; Score 1878; DB 10; Length 1255;
Best Local Similarity 83.0%; Pred. No. 3.5e-130;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

```
Qy 1 MELAALCRMGILLALLPPGAASOVCTGDMKLRLPASPTHLDMLRHL YOGCOVVOGNTL 60
|||
Db 1 MELAALCRMGILLALLPPGAASOVCTGDMKLRLPASPTHLDMLRHL YOGCOVVOGNTL 60
|||
Qy 61 ELTYLPTNASLSFLQDIQEVQGVYLAHNOVRLPQRLRIYRGTOLEFDNYALAVLDNG 120
|||
Db 61 ELTYLPTNASLSFLQDIQEVQGVYLAHNOVRLPQRLRIYRGTOLEFDNYALAVLDNG 120
|||
Qy 121 DPLNNTPTVYGASPGGLRELOLSLEILKGVLIQIRNPOLCYODITLWKDIFHKNNOLA 180
|||
Db 121 DPLNNTPTVYGASPGGLRELOLSLEILKGVLIQIRNPOLCYODITLWKDIFHKNNOLA 180
|||
Qy 181 LTLIDNRSRACHPCSPMKGSRGMESESDCOSLTRTYCAGGACARCKGPLPTDCCHEQC 240
|||
Db 181 LTLIDNRSRACHPCSPMKGSRGMESESDCOSLTRTYCAGGACARCKGPLPTDCCHEQC 240
|||
Qy 241 AACCTGPKHSDCLACLFHNSGICELHCPALVYNTDFESMPNPGRTYFGASCYTACP 300
|||
Db 241 AACCTGPKHSDCLACLFHNSGICELHCPALVYNTDFESMPNPGRTYFGASCYTACP 300
|||
Qy 301 YNLSLTDVSGCTLVCPPLHNOEYTAEDGTORCEKCSKPCARVCGYGLMEHLREVRATYSAN 360
|||
Db 301 YNLSLTDVSGCTLVCPPLHNOEYTAEDGTORCEKCSKPCARVCGYGLMEHLREVRATYSAN 360
|||
Qy 356 LRMOPG--PAHVPYSLFRPSMDLVSAFYSLPLAPLSPTSVPI-----SPYVGRGPD 405
|||
Db 356 LRMOPG--PAHVPYSLFRPSMDLVSAFYSLPLAPLSPTSVPI-----SPYVGRGPD 405
|||
Qy 406 --PDAAVAVNLSRYEG 419
|||
Db 406 --PDAAVAVNLSRYEG 419
|||
```



```

: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: HER500*-igM-CSF construct
US-09-821-883-4
```

```

Query Match          70.4%; Score 1610.5; DB 10; Length 697;
Best Local Similarity 75.4%; Pred. No. 8.1e-111;
Matches 319; Conservative 8; Mismatches 53; Indels 43; Gaps 7;
```

```

QY 16 LPPGASTVCTGTDKMLRLPASPEHIDMLRLHLYGCGVQVGNLELTYPNASTSFLQ 75
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 35 LARGASTVCTGTDKMLRLPASPEHIDMLRLHLYGCGVQVGNLELTYPNASTSFLQ 94
QY 76 DIOEVGYVLIANNQVROPQLRLRIVRGTOLEFEDNYALAVLDNGPLNNTTTPVGTASPG 135
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 95 DIOEVGYVLIANNQVROPQLRLRIVRGTOLEFEDNYALAVLDNGPLNNTTTPVGTASPG 154
QY 136 GLRELQLRSLTEILKGVLIQRNPQLCYODTILMKDIFHKNNQLALTLIDTNSRACHPC 195
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 155 GLRELQLRSLTEILKGVLIQRNPQLCYODTILMKDIFHKNNQLALTLIDTNSRACHPC 214
QY 196 SPCKSGSRMGESSEDCOSLTRVTCAGGACRCGRPLPTDCCHQCAAGCTGPRHSDCLAC 255
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 215 SPCKSGSRMGESSEDCOSLTRVTCAGGACRCGRPLPTDCCHQCAAGCTGPRHSDCLAC 274
QY 256 LPHNSGICELHCPALVTYNTDFEEMPNREGRTFGASCVTACPYNYLSTDVGSGTIVC 315
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 275 LPHNSGICELHCPALVTYNTDFEEMPNREGRTFGASCVTACPYNYLSTDVGSGTIVC 333
QY 316 PLHNOEVTADGTQRCCKSKPCARG-----THSLRPPAIV-VPLRMOPG----- 361
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 334 ---NFEKLAGGWNHRRHSSSTRSGGDLTLGLEPSEEBAPRSLAPSGAGSDVFDGD 390
QY 362 -----PAHVLSEFLRSPMLVSAFVSLAPL-SPTSVIISPVSGRGPD--- 405
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 391 LGWGAAGKGLSLPTDHPSPQLRYSED-----PVPPLPSETDGYVAPLTCSPQPEYVN 442
QY 406 -PD 407
   ||
DB 443 QPD 445
```

```

RESULT 14
US-09-854-356-8
: Sequence 8, Application US/09854356
: Patent No. US20020177567A1
: GENERAL INFORMATION:
: APPLICANT: Cheever, Martin A.
: APPLICANT: Gheysen, Dirk
: APPLICANT: Corixa Corporation
: APPLICANT: SmithKline Beecham Biologicals S. A.
: TITLE OF INVENTION: HER-2/neu Fusion Proteins
: FILE REFERENCE: 014058-009810PC
: CURRENT APPLICATION NUMBER: US/09/854,356
: PRIOR FILING DATE: 2001-05-09
: PRIOR APPLICATION NUMBER: US 09/493,480
: PRIOR FILING DATE: 2000-01-28
: PRIOR APPLICATION NUMBER: US 60/117,976
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 8
: LENGTH: 654
: TYPE: PRT
: ORGANISM: Rattus sp.
: FEATURE:
: OTHER INFORMATION: extracellular domain (ECD) of rat HER-2/neu
US-09-854-356-8
```

```

Query Match          70.3%; Score 1608.5; DB 9; Length 654;
Best Local Similarity 85.0%; Pred. No. 1.1e-110;
Matches 294; Conservative 15; Mismatches 36; Indels 1; Gaps 1;
```

```

QY 1 METALCRMGLLALLPPGASTVCTGTDKMLRLPASPEHIDMLRLHLYGCGVQVGNL 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 MELAAICRMGFLALLPPGASTVCTGTDKMLRLPASPEHIDMLRLHLYGCGVQVGNL 60
QY 61 ELTYLPTNASTSLFLDIOEVGYVLIANNQVROPQLRLRIVRGTOLEFEDNYALAVLDNG 120
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 61 ELTYVPAKASTSLFLDIOEVGYVLIANNQVROPQLRLRIVRGTOLEFEDNYALAVLDNR 120
QY 121 DPLNNTTPVT-GASPGGLELQLRSLTEILKGVLIQRNPQLCYODTILMKDIFHKNNQL 179
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 121 DPQDNVAASPTGTPREGLELQLRSLTEILKGVLIQRNPQLCYODTILMKDIFHKNNQL 180
QY 180 ALTLIDTNSRACHPCSPCKSGSRMGESSEDCOSLTRVTCAGGACRCGRPLPTDCCHQ 239
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 181 ARPYDITNSRACHPCAPRACKDNHMGESSEDCQILGTITGTCGACRCGRPLPTDCCHQ 240
QY 240 CAAGCTGPRHSDCLACLHNSGICELHCPALVTYNTDFEEMPNREGRTFGASCVTAC 299
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 241 CAAGCTGPRHSDCLACLHNSGICELHCPALVTYNTDFEEMPNREGRTFGASCVTTC 300
QY 300 PNYLSTDVGSGTIVCPPLHNOEVTADGTQRCCKSKPCARGTHSL 345
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 301 PNYLSTEVGSGTIVCPPNNOEVTADGTQRCCKSKPCARVCYGL 346
```

```

RESULT 15
US-09-854-356-2
: Sequence 2, Application US/09854356
: Patent No. US20020177567A1
: GENERAL INFORMATION:
: APPLICANT: Cheever, Martin A.
: APPLICANT: Gheysen, Dirk
: APPLICANT: Corixa Corporation
: APPLICANT: SmithKline Beecham Biologicals S. A.
: TITLE OF INVENTION: HER-2/neu Fusion Proteins
: FILE REFERENCE: 014058-009810PC
: CURRENT APPLICATION NUMBER: US/09/854,356
: PRIOR FILING DATE: 2001-05-09
: PRIOR APPLICATION NUMBER: US 09/493,480
: PRIOR FILING DATE: 2000-01-28
: PRIOR APPLICATION NUMBER: US 60/117,976
: PRIOR FILING DATE: 1999-01-29
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2
: LENGTH: 1256
: TYPE: PRT
: ORGANISM: Rattus sp.
: FEATURE:
: OTHER INFORMATION: rat HER-2/neu protein
: NAME/KEY: DOMAIN
: LOCATION: (1)..(654)
: OTHER INFORMATION: extracellular domain (ECD)
: NAME/KEY: DOMAIN
: LOCATION: (677)..(1256)
: OTHER INFORMATION: Intracellular domain (ICD)
: NAME/KEY: DOMAIN
: LOCATION: (721)..(998)
: OTHER INFORMATION: kinase domain (KD)
: NAME/KEY: DOMAIN
: LOCATION: (991)..(1256)
: OTHER INFORMATION: phosphorylation domain (PD)
: NAME/KEY: DOMAIN
: LOCATION: (991)..(1049)
: OTHER INFORMATION: fragment of the phosphorylation domain, preferred
: OTHER INFORMATION: portion (delta PD)
US-09-854-356-2
```

```

Query Match          70.3%; Score 1608.5; DB 9; Length 1256;
Best Local Similarity 85.0%; Pred. No. 2.3e-110;
Matches 294; Conservative 15; Mismatches 36; Indels 1; Gaps 1;
```

```
OY 1 MELALCRMGILLALPPGAASTOVCTGTDMLRLPASPEHLDMLRHLYOGCQVVOGNTL 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MELAMCRMGILLALPPGIAGTQVCTGTDMLRLPASPEHLDMLRHLYOGCQVVOGNTL 60
OY 61 ELTYLPTNASLSFLQDIOEVQGYVLAHNOVROYPLQRLRIVRGTQLFEDNVALAVLDNG 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 ELTYVPANASLSFLQDIOEVQGYMLAHNOVKRPVLPQRLRIVRGTQLFEDKVALAVLDNR 120
OY 121 DPLNNTTPTVT-GASPGGLRELOLRSLTEILKGGVLIQBNPOLCYODTILMKDIFHKNNOL 179
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 DPQDNVAASTPGKTPESGLRELOLRSLTEILKGGVLIQBNPOLCYQDNVLMKDVFRKNOL 180
OY 180 ALTLIDTNRSRACHPCSPMCKGSRGSESDQSLRTVCAGGACARCKGPLEPTDCHEQ 239
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 APVDIDTNRSRACPPCAPACKDNHCWESPEDCQILGTICTSGCARCKGRLEPTDCHEQ 240
OY 240 CAAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTAC 299
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 CAAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTFESMHNDEGRYTFGASCVTTC 300
OY 300 PNYVLSTDVGSCTLVCPPLHNOEYTAEDGTQRCCKGSKPCARGTHTSL 345
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 PNYVLSTEVGSCTLVCPPNNOEYTAEDGTQRCCKGSKPCARVCYGL 346
```

Search completed: March 4, 2003, 12:43:46
Job time : 20.6687 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:36:41 ; Search time 10.1526 Seconds
(without alignments)
1036.656 Million cell updates/sec

Title: US-09-234-208b-1

Perfect score: 79
Sequence: 1 GTHSLPRPAAVPYPLRMQP.....VGRGPDPAHVAVNLRYEG 79

Scoring table: OLIGO /
Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size : 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

A.Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/genesep-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/genesep-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/genesep-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	79	22	AAE09196 Human p68HER-2 ECD
2	79	100.0	83	22	AAE09185 Human p68HER-2 par
3	79	100.0	84	23	AAE20349 HER2 protein conta
4	73	92.4	79	22	AAE09195 Human p68HER-2 ECD
5	73	92.4	79	22	AAE09199 Human p68HER-2 ECD
6	72	91.1	79	22	AAE09188 Human p68HER-2 ECD
7	66	83.5	79	21	AAV97241 HER-2 C-terminal e
8	66	83.5	79	22	AAE09184 Human p68HER-2 ECD
9	66	83.5	79	22	AAE09186 Human p68HER-2 ECD
10	66	83.5	79	22	AAE09187 Human p68HER-2 ECD

11	66	83.5	79	23	AAE20350 Human HER2 Intron
12	57	72.2	79	22	AAE09194 Human p68HER-2 ECD
13	56	70.9	79	22	AAE09189 Human p68HER-2 ECD
14	55	69.6	79	22	AAE09197 Human p68HER-2 ECD
15	54	68.4	79	22	AAE09190 Human p68HER-2 ECD
16	51	64.6	79	22	AAE09191 Human p68HER-2 ECD
17	47	59.5	79	22	AAE09193 Human p68HER-2 ECD
18	41	51.9	79	22	AAE09198 Human p68HER-2 ECD
19	36	45.6	79	22	AAE09203 Human p68HER-2 ECD
20	17	21.5	79	21	AAV97239 HER-2 C-terminal p
21	17	21.5	79	22	AAE09180 Human p68HER-2 ECD
22	17	21.5	79	22	AAE09182 Human p68HER-2 ECD
23	17	21.5	79	23	AAE20347 Human HER2 Intron
24	15	19.0	419	22	AAE09212 Human p68HER-2 gen
25	15	19.0	419	22	AAE09213 Human p68HER-2 gen
26	15	19.0	419	22	AAE09216 Human p68HER-2 gen
27	14	17.7	419	22	AAE09181 Human p68HER-2 gen
28	14	17.7	419	22	AAE09200 Human p68HER-2 gen
29	14	17.7	419	22	AAE09203 Human p68HER-2 gen
30	14	17.7	419	22	AAE09204 Human p68HER-2 gen
31	14	17.7	419	22	AAE09205 Human p68HER-2 gen
32	14	17.7	419	22	AAE09206 Human p68HER-2 gen
33	14	17.7	419	22	AAE09207 Human p68HER-2 gen
34	14	17.7	419	22	AAE09208 Human p68HER-2 gen
35	14	17.7	419	22	AAE09209 Human p68HER-2 gen
36	14	17.7	419	22	AAE09210 Human p68HER-2 gen
37	14	17.7	419	22	AAE09211 Human p68HER-2 gen
38	14	17.7	419	23	AAE20348 Human truncated HE
39	14	17.7	420	21	AAV97240 Human p68HER-2, p
40	12	15.2	419	22	AAE09183 Human p68HER-2 gen
41	12	15.2	419	22	AAE09202 Human p68HER-2 gen
42	12	15.2	419	22	AAE09214 Human p68HER-2 gen
43	12	15.2	419	22	AAE09215 Human p68HER-2 gen
44	7	8.9	61	22	AAE61744 Human brain expres
45	7	8.9	61	22	AAE74540 Human bone marrow

ALIGNMENTS

RESULT 1
AAE09196
ID AAE09196 standard; peptide: 79 AA.
XX
AC AAE09196;
XX
DT 15-NOV-2001 (first entry)
XX
DE Human p68HER-2 ECDIIIA variant 11 encoded by HER-2 Intron 8.
XX
KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW Solid tumour; Cancer; polymorphism; cytostatic; gene therapy;
XX p68HER-2; ECDIIIA; variant.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 6 /note="p68HER-2 ECDIIIA (AAE09184) Pro substituted
FT FT with Leu"
FT FT
FT Misc-difference 73 /note="p68HER-2 ECDIIIA (AAE09184) Asp substituted
FT FT with Asn which is encoded by CAC"
XX
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US05327.
XX
XX 16-FEB-2000; 2000US-0506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.

```

XX Clinton G, Henner WD, Evans A;
PI
XX WPI: 2001-529934/58.
DR N-PSDB; AAD15865.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for
PI the treatment of hard tumors -
XX
XX Example 11: Page -: 61pp: English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Hereceptin, at an affinity of at least 10-8. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of Intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC the ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is p68HER-2 ECDIIIA variant
CC Note: The present sequence is not shown in the specification but is
CC derived from HER-2 Intron 8 encoded ECDIIIA sequence given in figure 8
CC (AAE09184).
XX
XX Sequence 79 AA:
SQ
XX
XX Query Match 100.0%; Score 79; DB 22; Length 79;
XX Best Local Similarity 100.0%; Pred. No. 2e-71;
XX Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 GTHSLPRPAAVPVRLMQPGPAHVPFLRPSWDLVSFAFYSILPLAPLSPTSPVPSV 60
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1 GTHSLPRPAAVPVRLMQPGPAHVPFLRPSWDLVSFAFYSILPLAPLSPTSPVPSV 60
OY 61 GRGPDPDAHVAVNLRSYEG 79
XX ||||||||||||||||||
DB 61 GRGPDPDAHVAVNLRSYEG 79
XX
XX RESULT 2
XX AAE09185 standard: Protein: 83 AA.
XX
XX AAE09185;
XX
XX 15-NOV-2001 (first entry)
XX
XX Human p68HER-2 partial protein containing ECDIIIA variant sequence.
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Hereceptin;
XX solid tumour; cancer; polymorphism; cytostatic; gene therapy;
XX p68HER-2; ECDIIIA; variant.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Domain 3..81
XX /label= ECDIIIA
XX /note= "Extracellular domain IIA of p68HER-2"
XX
XX Misc-difference 1 /note= "Encoded by CC"
XX
XX Misc-difference 8 /note= "Most commonly occurring ECDIIIA (AAE09184) Pro
XX substituted with Leu"
XX
XX Modified-site 75 /note= "Asn is N-glycosylated. Most commonly occurring
XX ECDIIIA (AAE09184) Asp substituted with Asn"
XX
XX

```

```

PN WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US05327.
XX
XX 16-FEB-2000; 2000US-0506079.
XX
XX (UWOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
PI
XX WPI: 2001-529934/58.
DR N-PSDB; AAD15854.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for
PI the treatment of hard tumors -
XX
XX Example 1: Fig 1: 61pp: English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Hereceptin, at an affinity of at least 10-8. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of Intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is human p68HER-2 partial protein
CC containing ECDIIIA variant sequence.
XX
XX Sequence 83 AA:
SQ
XX
XX Query Match 100.0%; Score 79; DB 22; Length 83;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-71;
XX Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 GTHSLPRPAAVPVRLMQPGPAHVPFLRPSWDLVSFAFYSILPLAPLSPTSPVPSV 60
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 3 GTHSLPRPAAVPVRLMQPGPAHVPFLRPSWDLVSFAFYSILPLAPLSPTSPVPSV 62
OY 61 GRGPDPDAHVAVNLRSYEG 79
XX ||||||||||||||||||
DB 63 GRGPDPDAHVAVNLRSYEG 81
XX
XX RESULT 3
XX AAE20349 standard: Protein: 84 AA.
XX
XX AAE20349;
XX
XX 18-JUN-2002 (first entry)
XX
XX HER2 protein containing extracellular domain (ECDIIIA).
XX
XX Endothelial growth factor receptor; EGFR; tumour; cytostatic; herstatin;
XX HER-2 receptor tyrosine kinase; squamous cell carcinoma; lung; colon;
XX glial cell tumour; cell growth.
XX
XX Unidentified.
XX
XX Key Location/Qualifiers
XX Misc-difference 1..2 /note= "Encoded by CCCGA"
XX
XX Modified-site 75..77 /note= "Asn is N-glycosylated"
XX
XX Misc-difference 81..82 /note= "Encoded by GGCTGAGACGGCCCTTCCGCCACCCACCCACAC
XX
XX

```

```

FT      TCCCTCAGTG"
FT      Misc-difference 83..84
FT      /note="Encoded by TGCT"
XX
XX      WO200214470-A2.
XX
XX      21-FEB-2002.
XX
XX      14-AUG-2001; 2001WO-US25502.
XX
XX      14-AUG-2000; 2000US-0638834.
XX
XX      (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX      Clinton GM;
XX
XX      WPI: 2002-269185/31.
XX
XX      N-PSDB; AAD32539.
XX
XX      Treating solid tumor characterized by expression of endothelial growth
XX      factor receptor, involves administering recombinant herstatin that
XX      binds to extracellular domain of the endothelial growth factor receptor
XX
XX      Example 11; Fig 1; 82pp; English.
XX
XX      The present invention relates to a method for treating a solid tumour
XX      characterised by endothelial growth factor receptor (EGFR) expression.
XX      The method involves administering an agent that binds to an extracellular
XX      domain (ECD) of EGFR. The invention also relates to a naturally occurring
XX      inhibitor of HER-2 receptor tyrosine kinase called herstatin. The co-
XX      expression of herstatin with p185HER2 causes a striking reduction in cell
XX      growth that corresponds with suppression of p185 autophosphorylation. The
XX      method or a pharmaceutical composition is useful for treating a solid
XX      tumour (selected from squamous cell carcinoma, lung carcinoma, colon
XX      carcinoma and glial cell tumour) characterised by EGFR expression. The
XX      present sequence is HER2 protein containing extracellular domain
XX      (ECDIIIA) sequence.
XX
XX      Sequence 84 AA:
SQ
Query Match      100.0%; Score 79; DB 23; Length 84;
Best Local Similarity 100.0%; Pred. No. 2.1e-71;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GTHSLPPRAAVPVPLRMQPGPAHPVLSFLRPSMDLVSAFYSLPLAPISPTSPVSPVSV 60
DB      3 GTHSLPPRAAVPVPLRMQPGPAHPVLSFLRPSMDLVSAFYSLPLAPISPTSPVSPVSV 62
QY      61 GRGPPDAHAVAVNLSRYEG 79
DB      63 GRGPPDAHAVAVNLSRYEG 81

RESULT 4
AAE09195
ID      AAE09195 standard; peptide: 79 AA.
XX
XX      AAE09195;
XX
XX      15-NOV-2001 (first entry)
XX
XX      Human p68HER-2 ECDIIIA variant 10 encoded by HER-2 intron 8.
XX
XX      HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX      solid tumour; cancer; polymorphism; cytostatic; gene therapy;
XX      p68HER-2; ECDIIIA; variant.
XX
XX      Homo sapiens.
XX
XX      Key      Location/Qualifiers
XX      Misc-difference 73 /note="p68HER-2 ECDIIIA (AAE09184) Asp substituted
FT

```

```

FT      with Asn which is encoded by CAC"
XX
XX      WO200161356-A1.
XX
XX      23-AUG-2001.
XX
XX      16-FEB-2001; 2001WO-US05327.
XX
XX      16-FEB-2000; 2000US-0506079.
XX
XX      (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX      Clinton G, Henner WD, Evans A;
XX
XX      WPI: 2001-529934/58.
XX
XX      N-PSDB; AAD15864.
XX
XX      New polypeptide, which binds to the extracellular domain of HER-2 for
XX      the treatment of hard tumors -
XX
XX      Example 11; Page -; 61pp; English.
XX
XX      The invention relates to novel HER-2 (herstatin-2) antagonist
XX      particularly a polypeptide that binds to the extracellular domain (ECD)
XX      of HER-2 at a site that is different from the binding site of humanised
XX      antibody, Herceptin, at an affinity of at least 1078. The present
XX      invention is based upon the initial discovery of an alternative HER-2
XX      mRNA transcript with 274 bp insert of intron 8. The translation product
XX      of the alternative transcript is a truncated HER-2 protein designated
XX      p68HER-2 which lacks the transmembrane and intracellular domains of
XX      p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX      The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX      the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX      nucleic acids encoding these are useful to treat, diagnose and identify
XX      solid tumours. The present sequence is p68HER-2 ECDIIIA variant
XX      encoded by polymorphic form of human HER-2 intron 8.
XX      Note: The present sequence is not shown in the specification but is
XX      derived from HER-2 intron 8 encoded ECDIIIA sequence given in figure 8
XX      (AAE09184).
XX
XX      Sequence 79 AA:
SQ
Query Match      92.4%; Score 73; DB 22; Length 79;
Best Local Similarity 100.0%; Pred. No. 2e-65;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      7 PRPAAVPVPLRMQPGPAHPVLSFLRPSMDLVSAFYSLPLAPISPTSPVSVGRGDP 66
DB      7 PRPAAVPVPLRMQPGPAHPVLSFLRPSMDLVSAFYSLPLAPISPTSPVSVGRGDP 66
QY      67 DAHVAVNLSRYEG 79
DB      67 DAHVAVNLSRYEG 79

RESULT 5
AAE09199
ID      AAE09199 standard; peptide: 79 AA.
XX
XX      AAE09199;
XX
XX      15-NOV-2001 (first entry)
XX
XX      Human p68HER-2 ECDIIIA variant 15 encoded by HER-2 intron 8.
XX
XX      HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX      solid tumour; cancer; polymorphism; cytostatic; gene therapy;
XX      p68HER-2; ECDIIIA; variant.
XX
XX      Homo sapiens.
XX
XX      Key      Location/Qualifiers
XX      Misc-difference 73
FT

```

```
FT /note= "p68HER-2 ECDIIIA (AAE09184) Asp substituted
FT with Asn"
XX
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US05327.
XX
XX 16-FEB-2000; 2000US-0506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
XX
XX WPI: 2001-529934/58.
XX
XX N-PSDB: AAD15868.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for
XX the treatment of hard tumors -
XX
XX Example 12: Page -: 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 1078. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is p68HER-2 ECDIIIA variant
XX encoded by polymorphic form of human HER-2 intron 8.
XX Note: The present sequence is not shown in the specification but is
XX derived from HER-2 intron 8 encoded ECDIIIA sequence given in figure 8
XX (AAE09184).
XX
XX Sequence 79 AA:
SQ
Query Match 92.4%; Score 73; DB 22; Length 79;
Best Local Similarity 100.0%; Pred. No. 2e-65;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 PRPAVPVPLRMQPGPAHPVLSFLRPSMDLSAFYSLPLAPLSPSPVSPVSGRGGPP 66
DB 7 PRPAVPVPLRMQPGPAHPVLSFLRPSMDLSAFYSLPLAPLSPSPVSPVSGRGGPP 66
QY 67 DAHVAVNLSRYEG 79
DB 67 DAHVAVNLSRYEG 79
RESULT 6
AAE09188
ID AAE09188 standard; peptide; 79 AA.
XX
XX AAE09188;
XX
XX 15-NOV-2001 (first entry)
XX
XX Human p68HER-2 ECDIIIA variant 3 encoded by HER-2 intron 8.
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX solid tumour; cancer; polymorphism; cytostatic; gene therapy;
XX p68HER-2; ECDIIIA; variant.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
```

```
FT Misc-difference 6
FT /note= "p68HER-2 ECDIIIA (AAE09184) Pro substituted
FT with Leu"
XX
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US05327.
XX
XX 16-FEB-2000; 2000US-0506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
XX
XX WPI: 2001-529934/58.
XX
XX N-PSDB: AAD15857.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for
XX the treatment of hard tumors -
XX
XX Example 11: Page -: 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 1078. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is p68HER-2 ECDIIIA variant
XX encoded by polymorphic form of human HER-2 intron 8.
XX Note: The present sequence is not shown in the specification but is
XX derived from HER-2 intron 8 encoded ECDIIIA sequence given in figure 8
XX (AAE09184).
XX
XX Sequence 79 AA:
SQ
Query Match 91.1%; Score 72; DB 22; Length 79;
Best Local Similarity 100.0%; Pred. No. 2e-64;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTHSLPRPAVPVPLRMQPGPAHPVLSFLRPSMDLSAFYSLPLAPLSPSPVSPVSV 60
DB 1 GTHSLPRPAVPVPLRMQPGPAHPVLSFLRPSMDLSAFYSLPLAPLSPSPVSPVSV 60
QY 61 GRGPPDAHVAV 72
DB 61 GRGPPDAHVAV 72
RESULT 7
AAV97241
ID AAV97241 standard; Protein; 79 AA.
XX
XX AAV97241;
XX
XX 04-DEC-2000 (first entry)
XX
XX HER-2 C-terminal extracellular domain IIIA.
XX
XX HER-2; erbB-2; oncogene; receptor-like tyrosine kinase; insertion;
XX extracellular domain IIIA; antagonist; intron 8; C-terminal extension;
XX truncated HER-2; p68; dimerization inhibitor; cytostatic.
XX
XX Homo sapiens.
XX
```


XX	16-FEB-2001; 2001WO-US05327.
PF	
XX	16-FEB-2000; 2000US-0506079.
XX	
PR	(UYOR-) UNIV OREGON HEALTH SCI.
PA	
XX	
PI	Clinton G, Henner WD, Evans A;
XX	
DR	WPI: 2001-529934/58.
XX	
DR	N-PSDB: AAD15855.
XX	
PT	New polypeptide, which binds to the extracellular domain of HER-2 for the treatment of hard tumors -
XX	
XX	Example 11: Page -, 61pp; English.
XX	
CC	The invention relates to novel HER-2 (herstatin-2) antagonist
CC	particularly a polypeptide that binds to the extracellular domain (ECD)
CC	of HER-2 at a site that is different from the binding site of humanised
CC	antibody, Herceptin, at an affinity of at least 10 ⁸ . The present
CC	invention is based upon the initial discovery of an alternative HER-2
CC	mRNA transcript with 274 bp insert of Intron 8. The translation product
CC	of the alternative transcript is a truncated HER-2 protein designated
CC	p68HER-2 which lacks the transmembrane and intracellular domains of
CC	p185HER-2 but contains ECD 1, II of the p185HER-2 and the novel ECDIIa.
CC	The ECDIIa-containing polypeptides bind tightly to, and thus antagonise
CC	the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC	nucleic acids encoding these are useful to treat, diagnose and identify
CC	solid tumours. The present sequence is p68HER-2 ECDIIa variant
CC	encoded by polymorphic form of human HER-2 intron 8.
CC	Note: The present sequence is not shown in the specification but is
CC	derived from HER-2 intron 8 encoded ECDIIa sequence given in figure 8
CC	(AAE09184).
XX	
XX	
SO	Sequence 79 AA:
XX	
Query Match	83.5%; Score 66; DB 22; Length 79;
Best Local Similarity	100.0%; Pred. No. 2e-58;
Matches	66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	7 PPRPAVPVPLRQPGPAHPVLSFLRPSMDLVSAFSLPLAPLSPGVSPTSGVGKGGPP 66
DB	7 PPRPAVPVPLRQPGPAHPVLSFLRPSMDLVSAFSLPLAPLSPGVSPTSGVGKGGPP 66
QY	67 DAHVAV 72
DB	67 DAHVAV 72
XX	
RESULT 10	
AAE09187	
ID	AAE09187 standard; peptide; 79 AA.
XX	
AC	AAE09187;
XX	
DT	15-NOV-2001 (first entry)
XX	
DE	Human p68HER-2 ECDIIa variant 2 encoded by HER-2 intron 8.
XX	
KW	HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW	solid tumour; cancer; polymorphism; cytostatic; gene therapy;
KW	p68HER-2; ECDIIa; variant.
XX	
OS	Homo sapiens.
XX	
PH	Key Location/Qualifiers
FT	Misc-difference 5
FT	/note="p68HER-2 ECDIIa (AAE09184) Leu substituted
FT	with Pro"
XX	
PN	WO200161356-A1.
XX	

[illegible]

PR	14-AUG-2000; 2000US--0638834.
XX	
PA	(UYOR-) UNIV OREGON HEALTH SCI.
XX	
PI	Clinton GM;
XX	
XX	WPI. 2002-269185/31.
DR	N-PSDB: AAD32540.
XX	
PT	Treating solid tumor characterized by expression of endothelial growth
PT	factor receptor, involves administering recombinant herstatin that
PT	binds to extracellular domain of the endothelial growth factor receptor
PT	-
PS	Example 11; Fig 8; 82pp; English.
XX	
CC	The present invention relates to a method for treating a solid tumour
CC	characterised by endothelial growth factor receptor (EGFR) expression.
CC	The method involves administering an agent that binds to an extracellular
CC	domain (ECD) of EGFR. The invention also relates to a naturally occurring
CC	inhibitor of HER-2 receptor tyrosine kinase called herstatin. The co-
CC	expression of herstatin with p185HER2 causes a striking reduction in cell
CC	growth that corresponds with suppression of p185 autophosphorylation. The
CC	method or a pharmaceutical composition is useful for treating a solid
CC	tumour (selected from squamous cell carcinoma, lung carcinoma, colon
CC	carcinoma and glial cell tumour) characterised by EGFR expression. The
CC	present sequence is a protein encoded by human HER2 intron 8.
CC	
SQ	Sequence 79 AA:
	Query Match 83.5%; Score 66; DB 23; Length 79;
	Best Local Similarity 100.0%; Pred. No. 2e-58;
	Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	7 PRRAPVPLRMQPPAPHPVLSFLRPSMDVSAFYSLPLAPLSPTSPVSPVSGGPP 66
DB	7 PRRAPVPLRMQPPAPHPVLSFLRPSMDVSAFYSLPLAPLSPTSPVSPVSGGPP 66
OY	67 DAHVAV 72
DB	67 DAHVAV 72
RESULT 12	
AAE09194	
ID	AAE09194 standard; peptide: 79 AA.
XX	
AC	AAE09194;
XX	
DT	15-NOV-2001 (first entry)
XX	
DE	Human p68HER-2 ECDIIIA variant 9 encoded by HER-2 intron 8.
XX	
KW	HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW	solid tumour; cancer; polymorphism; cytostatic; gene therapy;
XX	p68HER-2; ECDIIIA; variant.
XX	
OS	Homo sapiens.
XX	
XX	
EH	Key Location/Qualifiers
FT	Misc-difference 64
FT	/note="p68HER-2 ECDIIIA (AAE09184) Pro substituted
FT	with Leu"
XX	
PN	WO200161356-A1.
XX	
PD	23-AUG-2001.
XX	
PF	16-FEB-2001; 2001WO-US05327.
XX	
PR	16-FEB-2000; 2000US-0506079.
XX	
PA	(UYOR-) UNIV OREGON HEALTH SCI.

XX	Clinton G, Henner WD, Evans A;
PI	
XX	
DR	WPI: 2001-529934/58.
DR	N-PSDB; AAD15863.
XX	
PT	New polypeptide, which binds to the extracellular domain of HER-2 for
PT	the treatment of hard tumors -
XX	
PS	Example 11; Page -: 61pp; English.
XX	
CC	The invention relates to novel HER-2 (herstatin-2) antagonist
CC	particularly a polypeptide that binds to the extracellular domain (ECD)
CC	of HER-2 at a site that is different from the binding site of humanised
CC	antibody, Herceptin, at an affinity of at least 10 ⁸ . The present
CC	invention is based upon the initial discovery of an alternative HER-2
CC	mRNA transcript with 274 bp insert of intron 8. The translation product
CC	of the alternative transcript is a truncated HER-2 protein designated
CC	p68HER-2 which lacks the transmembrane and intracellular domains of
CC	p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
CC	The ECDIIIA-containing polypeptides bind tightly to an HER-2 ECD, and the
CC	HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC	nucleic acids encoding these are useful to treat, diagnose and identify
CC	solid tumours. The present sequence is p68HER-2 ECDIIIA variant
CC	encoded by polymorphic form of human HER-2 intron 8.
CC	Note: The present sequence is not shown in the specification but is
CC	derived from HER-2 intron 8 encoded ECDIIIA sequence given in figure 8
CC	(AAE09184).
XX	
SQ	Sequence 79 AA:
	Query Match 72.2%; Score 57; DB 22; Length 79;
	Best Local Similarity 100.0%; Pred. No. 2, 1e-49;
	Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY	7 PPAAPVPEPLRMQGPAPHPVLSFLRPSMDLVSAFTSLPLAPLSPISPVSVGSG 63
DB	7 PPAAPVPEPLRMQGPAPHPVLSFLRPSMDLVSAFTSLPLAPLSPISPVSVGSG 63
RESULT 13	
ID	AAE09189 standard; peptide: 79 AA.
AC	AAE09189;
XX	
DT	15-NOV-2001 (first entry)
XX	
DE	Human p68HER-2 ECDIIIA variant 4 encoded by HER-2 intron 8.
XX	
KW	HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW	Solid tumour; cancer; polymorphism; cytostatic; gene therapy;
KW	p68HER-2; ECDIIIA; variant.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 16
FT	/note="p68HER-2 ECDIIIA (AAE09184) leu substituted
FT	with Gln"
XX	
PN	WO200161356-A1.
XX	
PD	23-AUG-2001.
XX	
PF	16-FEB-2001; 2001WO-US05327.
XX	
PR	16-FEB-2000; 2000US-0506079.
XX	
PA	(UYOR-) UNIV OREGON HEALTH SCI.
XX	
PI	Clinton G, Henner WD, Evans A;
XX	

```
DR      WPI: 2001-529934/58.
DR      N-PSDB: AADI5858.
XX
XX      New polypeptide, which binds to the extracellular domain of HER-2 for
PT      the treatment of hard tumors -
PS
PS      Example 11: Page -: 61pp; English.
XX
CC      The invention relates to novel HER-2 (herstatin-2) antagonist
CC      particularly a polypeptide that binds to the extracellular domain (ECD)
CC      of HER-2 at a site that is different from the binding site of humanised
CC      antibody, Herceptin, at an affinity of at least 10-8. The present
CC      invention is based upon the initial discovery of an alternative HER-2
CC      mRNA transcript with 274 bp insert of intron 8. The translation product
CC      of the alternative transcript is a truncated HER-2 protein designated
CC      p68HER-2 which lacks the transmembrane and intracellular domains of
CC      p185HER-2 but contains ECD 1, 11 of the p185HER-2 and the novel ECDIIIA.
CC      The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC      the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC      nucleic acids encoding these are useful to treat, diagnose and identify
CC      solid tumours. The present sequence is p68HER-2 ECDIIIA variant
CC      encoded by polymorphic form of human HER-2 intron 8.
CC      Note: The present sequence is not shown in the specification but is
CC      derived from HER-2 intron 8 encoded ECDIIIA sequence given in figure 8
CC      (AAE09184).
CC      XX
SO      Sequence      79 AA:
SQ
Query Match          Best Local Similarity    70.9%; Score 56; DB 22; Length 79;
Matches   56; Conservative    100.0%; Pred. No. 2.1e-46;
                                         Mismatches 0; Indels 0; Gaps 0
OY      17 RMQGPAPHPVLSFLRPSMDVSAFYSLPLAPLSPTSVPISPVSVGGRPPDAHVAV 72
         |||||
Db       17 RMQGPAPHPVLSFLRPSMDVSAFYSLPLAPLSPTSVPISPVSVGGRPPDAHVAV 72
RESULT 14
AE03197 ID AAE09197 standard; peptide; 79 AA.
XX AC AAE09197;
XX DT 15-NOV-2001 (first entry)
XX DE Human p68HER-2 ECDIIIA variant 13 encoded by HER-2 intron 8.
XX KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;
KW p68HER-2; ECDIIIA; variant.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Misc-difference 17 /note="p68HER-2 ECDIIIA (AAE09184) Arg substituted
FT FT with Cys"
```

Pt	New polypeptide, which binds to the extracellular domain of HER-2 for the treatment of hard tumors -
Pt	
xx	
PS	Example 12; Page -: 6lpp; English.
CC	The invention relates to novel HER-2 (herstatin-2) antagonist particularly a polypeptide that binds to the extracellular domain (ECD) of HER-2 at a site that is different from the binding site of humanised antibody, Herceptin, at an affinity of at least 10 ⁻⁸ . The present invention is based upon the initial discovery of an alternative HER-2 mRNA transcript with 274 bp insert of intron 8. The translation product of the alternative transcript is a truncated HER-2 protein designated p68HER-2 which lacks the transmembrane and intracellular domains of pl5HER-2 but contains ECD I, II of the pl85HER-2 and the novel ECDIIia. The ECDIIia-containing polypeptides bind tightly to, and thus antagonise the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the nucleic acids encoding these are useful to treat, diagnose and identify solid tumours. The present sequence is p68HER-2 intron 8.
CC	Note: The present sequence is not shown in the specification but is derived from HER-2 intron 8 encoded ECDIIia sequence given in figure 8 (AAE09184).
CC	
SQ	Sequence 79 AA;
Oy	Query Match 69.6%; Score 55; DB 22; Length 79; Best Local Similarity 100.0%; Pred. No. 2.le-47; Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Dd	18 MOGGAHPPLSLRSMDLVSAFYLSLPLALPSTVPISPVYSGRPDDAHAYV 72 18 MQPGAHPVLPLSRSWDLVSFAFYSLLPLAPlSPTSVIPSPVASVGKPPDPAHVAV 72
ID	AEE09190 standard; peptide: 79 AA. AE09190
AC	AAE09190;
XX	
DT	15-NOV-2001 (first entry)
DE	Human p68HER-2 ECDIIia variant 5 encoded by HER-2 intron 8.
KX	HER-2; herstatin; antagonist; extracelular domain; ECD; Herceptin;
KW	solid tumour; cancer; polymorphism; cytostatic; gene therapy;
KM	p68HER-2; ECDIIia; variant.
OS	Homo sapiens.
FH	Key Location/Qualifiers
FT	Misc-difference 18 /note= "p68HER-2 ECDIIia (AAE09184) Met substituted with Leu which is encoded by ATA"
PN	WO200161356-A1.
PD	23-AUG-2001.
PF	16-FEB-2001; 2001WO-USO5327.
PR	16-FEB-2000; 2000US-0506079.
PA	(UTOR-) UNIV OREGON HEALTH SCI.
PI	Clinton G, Henner WD, Evans A;
DR	MPJ: 2001-529934/58.
N-PDB:	AAD15859.
Pt	New polypeptide, which binds to the extracellular domain of HER-2 for the treatment of hard tumors -
xx	

PS Example 11; Page -; 61pp; English.
xy

CC The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 10⁸. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p85HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD, II of the p185HER-2 and the novel ECDIIA.
CC The ECDIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is p85HER-2 ECDIIA variant
CC encoded by polymorphic form of human HER-2 intron 8.
CC Note: The present sequence is not shown in the specification but is
CC derived from HER-2 intron 8 encoded ECDIIA sequence given in figure 8
CC (AAE09184).

Query Match	68.4%;	Score 54;	DB 22;	Length 79;
Best Local Similarity	100.0%;	Pred. No. 2.1e-46;		
Matches 54; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

Qy 19 QPRAHPVLSTFLRPSMDLSAFYSLPLAPLSPTSVIPISVSVGRRPPDAHAAV 72
Db 19 QPRAHPVLSTFLRPSMDLSAFYSLPLAPLSPTSVIPISVSVGRRPPDAHAAV 72

Search completed: March 4, 2003, 12:44:53
Job time : 10.1526 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:43:27 ; Search time 8.88353 Seconds

(without alignments)
854,910 Million cell updates/sec

Title: US-09-234-208b-1

Perfect score: 79

Sequence: 1 GTHSLPRPAAVPPLRMQP.....VGRGPDPAHVAVNLSEYEG 79

Scoring table: OLIGO

Searched: 283224 seqs, 96134422 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	10.1	520	2 C70311	hypothetical prote
2	8	10.1	865	2 A85032	hypothetical prote
3	8	10.1	879	2 S49910	chloroplast outer
4	8	10.1	1503	2 T01098	chloroplast outer
5	7	8.9	121	2 AH1848	hypothetical prote
6	7	8.9	189	2 S04670	hypothetical prote
7	7	8.9	207	2 A69941	hypothetical prote
8	7	8.9	269	1 D43706	capsular polysach
9	7	8.9	292	2 S40979	serine O-acetyltra
10	7	8.9	328	2 G84826	hypothetical prote
11	7	8.9	382	2 B88561	hypothetical prote
12	7	8.9	389	2 A97577	protein F58A4_7b f
13	7	8.9	389	2 AH2797	hypothetical prote
14	7	8.9	397	2 C83470	conserved hypotet
15	7	8.9	436	2 B70473	hypothetical prote
16	7	8.9	464	2 AF2180	protoporphyrinogen
17	7	8.9	495	2 T17478	dianonopimelate de
18	7	8.9	555	2 A26639	hypothetical prote
19	7	8.9	574	2 S01272	regulatory protein
20	7	8.9	595	2 T15862	regulatory protein
21	7	8.9	840	2 T36175	hypothetical prote
22	7	8.9	895	2 A45554	probable large ATP
23	7	8.9	906	1 IJM5CN	rhophy complex po
24	7	8.9	2240	2 T37057	N-cadherin precurs
25	6	7.6	59	2 T07432	probable multi-dom
26	6	7.6	65	2 C72558	photosystem II pro
27	6	7.6	74	2 F82764	hypothetical prote
28	6	7.6	94	2 H81190	hypothetical prote
29	6	7.6	98	2 S68136	NADH2 dehydrogenas

30	6	7.6	101	2 H72493	hypothetical prote
31	6	7.6	119	2 S75551	hypothetical prote
32	6	7.6	120	2 D83166	hypothetical prote
33	6	7.6	122	2 AB2199	hypothetical prote
34	6	7.6	126	2 S25104	hypothetical prote
35	6	7.6	128	2 G71264	MSF14 protein - ma
36	6	7.6	128	2 A87510	conserved hypotet
37	6	7.6	129	2 B83173	hypothetical prote
38	6	7.6	142	2 C58723	hypothetical prote
39	6	7.6	142	2 E72571	hypothetical prote
40	6	7.6	145	1 J00347	hypothetical prote
41	6	7.6	150	2 A69289	capsid protein VP3
42	6	7.6	159	2 A69289	conserved hypotet
43	6	7.6	163	2 AB2025	conserved hypotet
44	6	7.6	171	2 T43959	hypothetical prote
45	6	7.6	182	2 S76346	hypothetical prote

ALIGNMENTS

RESULT 1
C70311
hypothetical protein aq_116 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jun-2000
C:Accession: C70311
R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300, PMID:98196666; PMID:9537320
A:Accession: C70311
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-520 <AOF>
A:Cross-references: GB:AE000674; NID:g2982850; PIDN:AA06482.1; PID:g2982861; GB:AE00
A:Experimental source: strain VF5
C:Genetics:
A:Gene: aq_116
C:Superfamily: Aquifex aeolicus hypothetical protein aq_116
Query Match 10.1%; Score 8; DB 2; Length 520;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 47 PLSPSTVP 54
DB 2 PLSPSTVP 9
RESULT 2
A85032
hypothetical protein AT4902510 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: A85032
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold SP
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; PMID:20083488; PMID:10617198
A:Accession: A85032
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-865 <STO>
A:Cross-references: GB:NC_001268; NID:g7269011; PIDN:CAB80744.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4902510
A:Map position: 4
Query Match 10.1%; Score 8; DB 2; Length 865;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 PAAVPVPL 16
|||||||
Db 584 PAAVPVPL 591

RESULT 3
S49910
chloroplast outer envelope protein OEP86 precursor - garden pea
N:Alternate names: chloroplast import-associated protein IAR86, GTP-binding
C:Species: Pisum sativum (garden pea)
C:Date: 26-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 02-Feb-2001
C:Accession: S49910; A55386; A55171
R:Hirsch, S.; Soll, J.
submitted to the EMBL Data Library, March 1994
A:Reference number: S49910
A:Accession: S49910

A:Molecule type: mRNA
A:Residues: 1-879 <HIR>
A:Cross-references: EMBL:231581; NID:959957; PID:959958
R:Hirsch, S.; Muckel, E.; Heemeyer, F.; von Heijne, G.; Soll, J.
Science 266, 1989-1992, 1994
A:Title: A receptor component of the chloroplast protein translocation machinery.
A:Reference number: A55386; MUID:9509324; PMID:7801125
A:Accession: A55386
A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA
A:Residues: 1-137, 'O', 139-875, 877-879 <HI2>
R:Kessler, F.; Blobel, G.; Patel, H.A.; Schnell, D.J.
Science 266, 1035-1039, 1994
A:Title: Identification of two GTP-binding proteins in the chloroplast protein import ma
A:Reference number: A55171; MUID:95063938; PMID:7973656
A:Accession: A55171
A:Molecule type: mRNA
A:Residues: 1-372, 'E00', 376-879 <KE5>
A:Cross-references: GB:U36657; NID:9576508; PIDN:AA53276.1; PID:9576509
C:Keywords: chloroplast; GTP binding; membrane protein; nucleotide binding; P-loop
F:245-252/Region: nucleotide-binding motif A (P-loop)

Query Match 10.1%; Score 8; DB 2; Length 879;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 PAAVPVPL 16
|||||||
Db 599 PAAVPVPL 606

RESULT 4
T01098
chloroplast outer envelope protein OEP86 homolog T10P11.19 - Arabidopsis thaliana
N:Alternate names: chloroplast import-associated protein, GTP-binding; protein T14P8.24
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 02-Feb-2001
C:Accession: T01098; T01299
R:Kaplan, N.; Johnson, D.; Schut, K.; Gnoj, L.; Hoffman, J.; Tili, S.; de la Bastide, M.
hi, M.; Martienssen, R.; Chen, E.Y.; Wilson, R.; McCombie, W.R.
submitted to the EMBL Data Library, November 1998
A:Description: Sequence of A. thaliana BAC T10P11 from chromosome IV.
A:Reference number: Z14248
A:Accession: T01098
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1503 <KAP>
A:Cross-references: EMBL:AC002330; NID:92262135; PID:93892053
A:Experimental source: cultivar Columbia
R:Kallick, J.; Elliott, G.; Cloud, J.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of A. thaliana T14P8.
A:Reference number: Z14290
A:Accession: T01299
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-1503 <KAL>
A:Cross-references: EMBL:AF069298; NID:93193282; PID:93193301
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 4
A:Introns: 22/1
A:Note: T10P11.19; T14P8.24
C:Keywords: chloroplast; GTP binding; membrane protein; nucleotide binding; P-loop
F:862-869/Region: nucleotide-binding motif A (P-loop)

Query Match 10.1%; Score 8; DB 2; Length 1503;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 PAAVPVPL 16
|||||||
Db 1222 PAAVPVPL 1229

RESULT 5
AH1848
hypothetical protein all0337 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AH1848
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kurlitz, T.; Sasamoto, S.; Matanabe, A.; Iri
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 6, 205-213, 2001
A:Title: Complete genomic sequence of the filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AH1807; MUID:21595285; PMID:11759840
A:Accession: AH1848
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-121 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA072295.1; PID:917129682; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all0337

Query Match 8.9%; Score 7; DB 2; Length 121;
Best Local Similarity 100.0%; Pred. No. 8.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 PVLSPFLR 31
|||||||
Db 56 PVLSPFLR 62

RESULT 6
S04670
hypothetical protein 5 - Rhodospseudomonas blastic
C:Species: Rhodospseudomonas blastic
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 18-Jun-1993
C:Accession: S04670
R:Tybuliewicz, V.L.J.; Falk, G.; Walker, J.E.
J. Mol. Biol. 179, 185-214, 1984
A:Title: Rhodospseudomonas blastica atp operon. Nucleotide sequence and transcription.
A:Reference number: S04666; MUID:85058188; PMID:6209404
A:Accession: S04670
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-189 <TYB>

Query Match 8.9%; Score 7; DB 2; Length 189;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 50 PTVSPIS 56
|||||||
Db 84 PTVSPIS 90

```
RESULT 7
A:Accession: A69941
capsular polysaccharide homolog yypq - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C:Accession: A69941
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillette, S.; Brusch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
lech, J.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koeter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinis,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadie, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeder, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serx
auech, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: A69941
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-207 <KUN>
A:Cross-references: GB:299115; GB:AL009126; NID:92634478; PIDN:CAB14085.1; PID:ell83614;
A:Experimental source: strain 168
C:Genetics:
A:Gene: yypq

Query Match
Best Local Similarity 100.0%; Score 7; DB 2; Length 207;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LPRPAV 12
| | | | |
DB 125 LPRPAV 131

RESULT 8
D:3706
serine O-acetyltransferase (EC 2.3.1.30) nlf - Azotobacter chroococcum
C:Species: Azotobacter chroococcum
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 05-May-2000
C:Accession: D:3706
R:Evans, D.J.; Jones, R.; Woodley, P.R.; Wilborn, J.R.; Robson, R.L.
J. Bacteriol. 173, 5457-5469, 1991
A:Title: Nucleotide sequence and genetic analysis of the Azotobacter chroococcum nlfusw
A:Reference number: A43706; MUID:91358323; PMID:1885524
A:Accession: D:3706
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-269 <EVA>
A:Cross-references: EMBL:M60090; NID:q142386; PIDN:AAA22162.1; PID:q142390
C:Superfamily: Bacillus serine acetyltransferase; serine acetyltransferase homology
C:Keywords: acyltransferase; coenzyme A
F:10-169/Domain: serine acetyltransferase homology <SAT>

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 269;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LPRPAV 12
| | | | |
DB 52 LPRPAV 58

RESULT 9
S:40979
hypochemical protein F58A.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Feb-1997
```

```
C:Accession: S:40979
R:Berk, M.
submitted to the EMBL Data Library, February 1992
A:Reference number: S:40973
A:Accession: S:40979
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-292 <BER>
A:Cross-references: EMBL:D22179
C:Genetics:
A:Insertions: 60/2; 133/1; 273/1
C:Keywords: nucleus

Query Match
Best Local Similarity 100.0%; Score 7; DB 2; Length 292;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 APLSPTS 52
| | | | |
DB 48 APLSPTS 54

RESULT 10
G:4826
hypochemical protein At2g40220 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: G:4826
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617157
A:Accession: G:4826
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-328 <STO>
A:Cross-references: GB:AE002093; NID:96598941; PIDN:AAF18736.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g40220
A:Map position: 2

Query Match
Best Local Similarity 100.0%; Score 7; DB 2; Length 328;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LPRPAA 11
| | | | |
DB 144 LPRPAA 150

RESULT 11
B:8561
protein F58A.7b [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: B:8561
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C-
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A:Accession: B:8561
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-382 <STO>
A:Cross-references: GB:chr_III; PIDN:CAA80170.1; PID:q3877829; GSPDB:GN00021; CESP:F5
C:Genetics:
A:Gene: F58A.7b
A:Map position: 3
```

Query Match 8.9%; Score 7; DB 2; Length 382;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 46 APLSPTS 52
|||||||
DB 48 APLSPTS 54

RESULT 12

A97577
hypothetical protein AGR_C3311 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: A97577
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; PMID:11743194
A:Accession: A97577
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-389 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK87570.1; PID:G15156910; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_3311
A:Map position: circular chromosome

Query Match 8.9%; Score 7; DB 2; Length 389;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 PAAPVP 15
|||||||
DB 160 PAAPVP 166

RESULT 13

AH2797
conserved hypothetical protein Atu1800 [imported] - Agrobacterium tumefaciens (strain C5
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AH2797
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClie
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AH2797
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-389 <KUR>
A:Cross-references: GB:AE008688; PIDN:AAL42798.1; PID:G17740243; GSPDB:GN00186
C:Genetics:
A:Gene: Atu1800
A:Map position: circular chromosome

Query Match 8.9%; Score 7; DB 2; Length 389;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 PAAPVP 15
|||||||
DB 160 PAAPVP 166

RESULT 14

C83470

hypothetical protein PA1412 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: C83470

R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: C83470

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-397 <STO>
A:Cross-references: GB:AE004570; GB:AE004091; NID:9947350; PIDN:AA04801.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA1412

Query Match 8.9%; Score 7; DB 2; Length 397;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 LLRPAA 11
|||||||
DB 187 LLRPAA 193

RESULT 15

B70473
protoporphyrinogen oxidase - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: B70473
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: B70473
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-436 <AGP>
A:Cross-references: GB:AE000768; NID:92984249; PIDN:AAC07778.1; PID:G2984251; GB:AE00
A:Experimental source: strain VP5
C:Genetics:
A:Gene: hemc

Query Match 8.9%; Score 7; DB 2; Length 436;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 26 VLSFLRP 32
|||||||
DB 428 VLSFLRP 434

Search completed: March 4, 2003, 12:47:41
Job time : 10.8835 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:37:16 ; Search time 5.39357 Seconds
(without alignments)
607.506 Million cell updates/sec

Title: US-09-234-208b-1

Perfect score: 79

Sequence: 1 GTHSLPRAAVPVPLRMQP.....VGRCPDPAVAVNLRYEG 79

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	8.9	114	GON2_TUPGB	095336 tupala glis
2	7	8.9	189	YAN5_RHOBL	P05448 rhodopsendo
3	7	8.9	207	YQGP_BACSU	P54183 bacillus su
4	7	8.9	269	NIFP_AZCH	P23145 azotobacter
5	7	8.9	380	YMH7_CAEEL	P34474 caenorhabdi
6	7	8.9	499	CPM1_ONCMY	092088 oncorhynch
7	7	8.9	574	ZEST_DROME	P09556 drosophila
8	7	8.9	635	SL56_HUMAN	092889 homo sapien
9	7	8.9	677	GCP3_MOUSE	P58854 mus musculi
10	7	8.9	906	CAD2_MOUSE	P15116 mus musculi
11	7	8.9	907	GCP3_HUMAN	0966W5 homo sapien
12	7	8.9	968	ST10_HUMAN	094804 homo sapien
13	6	7.6	56	PSBK_PINTH	P41598 pinus thund
14	6	7.6	111	YCX5_OENHO	09amc3 oenothera h
15	6	7.6	126	ML14_MALIZ	001900 zea mays (m
16	6	7.6	141	RL9_HABCO	002376 haemochus
17	6	7.6	145	VP3_BPCHP	P19194 bacterioph
18	6	7.6	159	YXHD_BACSU	P46328 bacillus su
19	6	7.6	195	GSP1_ECOLI	P45761 escherichia
20	6	7.6	198	LMBV_CHICK	001636 gallus galli
21	6	7.6	204	LAH4_MOUSE	P51827 mus musculi
22	6	7.6	205	ESMC_DROME	001070 drosophila
23	6	7.6	206	COX3_BACFI	004442 bacillus fi
24	6	7.6	209	ADPE_ECOLI	P36651 escherichia
25	6	7.6	214	VATD_DESSY	006506 desulfuroco
26	6	7.6	215	CRB6_ODOSI	P49488 odontella s
27	6	7.6	220	ACPD_STRCO	095196 streptomyce
28	6	7.6	220	BIOD_AQUAE	006832 aquifex aeo
29	6	7.6	235	GCH1_SCHPO	013774 schizosacch
30	6	7.6	239	PRR4_BOVIN	P18917 bos taurus
31	6	7.6	241	LAT_RAT	070601 ratius norv
32	6	7.6	242	LAT_MOUSE	054957 mus musculi
33	6	7.6	248	PT16_LYCCE	004682 lycopersico

34	6	7.6	250	1	MTR3_YEAST	P48240 saccharomyc
35	6	7.6	251	1	Y4EP_RHISN	P55429 rhizobium s
36	6	7.6	260	1	YH20_PASMU	09ck45 pasteurella
37	6	7.6	263	1	COO4_CAEEL	P91428 caenorhabdi
38	6	7.6	290	1	AR11_MESAU	P50292 mesocricetu
39	6	7.6	290	1	MPN_HUMAN	09bgr3 homo sapien
40	6	7.6	295	1	SSO2_YEAST	P39926 saccharomyc
41	6	7.6	301	1	MPEV_SYNPY	002178 synchococc
42	6	7.6	301	1	VF34_SYNP3	P74220 synchocyst
43	6	7.6	303	1	RDCG_ECO57	08xcal escherichia
44	6	7.6	303	1	RDCG_ECOLI	P36767 escherichia
45	6	7.6	303	1	RDCG_SALTY	08zre6 salmoneella

ALIGNMENTS

```

RESULT 1
ID GON2_TUPGB          STANDARD:      PRT: 114 AA.
AC 095336:
DT 15-DEC-1998 (rel. 37, Created)
DT 15-DEC-1998 (rel. 37, Last sequence update)
DT 30-MAY-2000 (rel. 39, Last annotation update)
DE Progonadolibertin II precursor [Contains: Gonadolibertin II (LHRH II)
DE (luteinizing hormone releasing hormone II) (Gonadotropin releasing
DE hormone II) (GNRH II) (Luliberin II); Gnrh-associated peptide II].
GN GNRH2.
OS Tupala glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupata.
OX NCBI_TaxID=9396;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RX MEDLINE=97079639; PubMed=8921350;
RA Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
RA Fernald R.D.;
RT "Characterization of two new preproGNRH mRNAs in the tree shrew:
RT first direct evidence for mesencephalic Gnrh gene expression in a
RT placental mammal.";
RT Gen. Comp. Endocrinol. 104:7-19(1996).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: MIDBRAIN.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U63327; AAB16838.1; -.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Signal.
FT SIGNAL 1 25
FT CHAIN 26 114 PROGNADOLIBERTIN II.
FT PEPTIDE 26 35 GONADOLIBERTIN II.
FT PEPTIDE 39 114 GNRH-ASSOCIATED PEPTIDE II.
FT MOD_RES 26 PYRROLIDONE CARBOXYLIC ACID
FT MOD_RES 26 (BY SIMILARITY).
FT MOD_RES 35 AMIDATION (G-36 PROVIDE AMIDE GROUP).
SQ SEQUENCE 114 AA: 12123 MW: 680E90E1C6869EBC1 CXC64;
Query Match 8.9%; Score 7; DB 1; Length 114;
Best Local Similarity 100.0%; Pred. No. 4.2;

```

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 PRPAVP 13
 |||||
 Db 105 PRPAVP 111

RESULT 2

YAF5_RHOBL STANDARD: PRT: 189 AA.
 ID YAF5_RHOBL
 AC P05448;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 01-NOV-1988 (Rel. 09, Last annotation update)
 DE ATP synthase subunits region ORF 5.
 OS Rhodopseudomonas blattica.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
 CC Rhodobacter.
 OX NCBI_TaxID=1075;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85058188; PubMed=6209404;
 RA Tyduniewicz V.L.J., Falk G., Walker J.E.;
 RT "Rhodopseudomonas blattica atp operon. Nucleotide sequence and
 transcription.";
 RL J. Mol. Biol. 179:185-214(1984).

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: 200018; CAA77311.1; -
 DR PIR: S04670; S04670.
 KW Hypothetical protein.
 SQ SEQUENCE 189 AA; 20759 MW; 68756B3A3C8C466 CRC64;

Query Match 8.9%; Score 7; DB 1; Length 189;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 50 PRSPVIS 56
 |||||
 Db 84 PRSPVIS 90

RESULT 3
 YPOP_BACSU STANDARD: PRT: 207 AA.
 ID YPOP_BACSU
 AC P54183;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein ypop (Fragment).
 GN YPOP.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / Marburg;
 RX MEDLINE=96349105; PubMed=8760912;
 RA Sorokin A.V., Azevedo V., Zumbstein E., Galleron N., Ehrlich S.D.,
 RA Serior P.;
 RT "Sequence analysis of the Bacillus subtilis chromosome region between
 RT the sera and kgd loci cloned in a yeast artificial chromosome.";
 RL Microbiology 142:2005-2016(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

 CC EMBL: L77246; AAA96649.1; -
 DR EMBL: 299115; CAB14085.1; -
 DR Subtilist; BG11629; YPOP.
 DR Interpro: IPR003869; Polysac_Capp.
 DR Pfam: PF02719; Polysac synt_2; 1.
 KW Hypothetical protein; Complete proteome.
 FT NON_TER
 SQ SEQUENCE 207 AA; 23286 MW; F3EA64945DE249CD CRC64;

Query Match 8.9%; Score 7; DB 1; Length 207;
 Best Local Similarity 100.0%; Pred. No. 7.2;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LPPRAVP 12
 |||||
 Db 125 LPPRAVP 131

RESULT 4

NIFP_AZCOH STANDARD: PRT: 269 AA.
 ID NIFP_AZCOH
 AC P23145;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE Probable serine acetyltransferase (EC 2.3.1.30) (SAT).
 GN NIFP.
 OS Azotobacter chroococcum mcd 1.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Azotobacter.


```

OX NCBI_TaxID=355;
RN
  SEQUENCE FROM N.A.
RX MEDLINE=91358323; PubMed=1885524;
RA Evans D.J., Jones R., Woodley P.R., Wilborn J.R., Robson R.L.;
RT "Nucleotide sequence and genetic analysis of the Azotobacter
   chroococcum nifUSVWZM gene cluster, including a new gene (nifP) which
   encodes a serine acetyltransferase."
RL J. Bacteriol. 173:5457-5469(1991).
CC
  -i- FUNCTION: PROBABLE SERINE ACETYLTRANSFERASE REQUIRED FOR
   OPTIMIZING THE EXPRESSION OF NITROGENASE ACTIVITY. NIFP MAY BE
   REQUIRED TO BOOST RATES OF SYNTHESIS OR INTRACELLULAR
   CONCENTRATIONS OF CYSTEINE OR METHIONINE.
CC
  -i- CATALYTIC ACTIVITY: Acetyl-CoA + L-serine = CoA + O-acetyl-L-
   serine.
CC
  -i- SIMILARITY: BELONGS TO THE CYSE/LACA/LPYA/NOXL FAMILY OF
   ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).
CC
  -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (see http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC
  -----
DR EMBL: M60090; AAA22162.1; -.
DR PIR: DA3706; DA3706.
DR InterPro: IPR001451; Hexapep_transf.
DR Pfam: PF00132; hexapep; 4.
DR TIGRfams: TIGR01172; cyse; 1.
DR PROSITE: PS00101; HEXAPEP_TRANSFERASES; 1.
KW Nitrogen fixation; Transferase; Acyltransferase; Repeat;
   Cysteine biosynthesis.
SQ
  SEQUENCE 269 AA; 28578 MW; E0BBCC982B666DBB CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 269;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LPRPAV 12
   |||||
Db 52 LPRPAV 58

RESULT 5
YMH7_CAEEL STANDARD: PRT; 380 AA.
ID YMH7_CAEEL STANDARD: PRT; 380 AA.
AC P34474;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical 31.8 kDa protein F58A4.7 in chromosome III.
GN F58A4.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Chordata; Rhabdillida; Rhabdillidae;
   Rhabdillidae; Pelodidae; Caenorhabditis.
CC
  NCBI_TaxID=6239;
CX [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
   Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
   Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
   Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
   Lathille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
   Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,
   Sims M., Smailson N., Smith A., Smith M., Sonhammer E., Staden R.,
   Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
   Waterston R., Watson A., Weinstock L., Wilkinson-Spratt J.,
   Woldman P.;

```

```

RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
   elegans."
RL Nature 368:32-38(1994).
CC
  -i- SUBCELLULAR LOCATION: Nuclear (Potential).
CC
  -i- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
   TRANSCRIPTION FACTORS. STRONGEST TO TRANSCRIPTION FACTOR AP-4.
CC
  -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (see http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC
  -----
DR EMBL: 222179; CA860167.1; -.
DR PIR: S40979; S40979.
DR WormPep: F58A4.7; CB01018.
DR InterPro: IPR001092; HLH_basic.
DR Pfam: PF00010; HLH; 1.
DR SMART: SM00353; HLH; 1.
DR PROSITE: PS00038; HLH_1; 1.
DR PROSITE: PS50888; HLH_2; 1.
KW Hypothetical protein; Transcription regulation; DNA-binding;
   Nuclear protein.
FT DNA_BIND 61 73 BASIC DOMAIN.
FT DOMAIN 74 113 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
SQ
  SEQUENCE 380 AA; 40519 MW; EE2B52D1EB3EFD33 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 380;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 APLSPS 52
   |||||
Db 48 APLSPS 54

RESULT 6
CPM1_ONCMY STANDARD: PRT; 499 AA.
ID CPM1_ONCMY STANDARD: PRT; 499 AA.
AC Q92088;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytochrome P450 2M1 (EC 1.14.14.1) (CYPM1) (Lauric acid omega-6-
   hydroxylase) (LMC1).
GN CYP2M1.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
CX [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Shasta; TISSUE=Liver;
RX MEDLINE=98248987; PubMed=9587416;
RA Yang Y.-H., Wang J.-L., Miranda C.L., Buhler D.R.;
RT "CYP2M1: Cloning, sequencing, and expression of a new cytochrome P450
   from rainbow trout liver with fatty acid (omega-6)-hydroxylation
   activity."
RL Arch. Biochem. Biophys. 352:271-280(1998).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=90386619; PubMed=2403346;
RA Miranda C.L., Wang J.L., Henderson M.C., Williams D.E., Buhler D.R.;
RT "Regiospecificity in the hydroxylation of lauric acid by rainbow trout
   hepatic cytochrome P450 isozyms."
RL Biochem. Biophys. Res. Commun. 171:537-542(1990).
CC
  -i- FUNCTION: HAS (OMEGA-6)-HYDROXYLATION ACTIVITY TOWARD LAURIC ACID.
   -i- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
   oxidized flavoprotein + H(2)O.
CC
  -i- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum (By

```

CC similarity).
CC -1- TISSUE SPECIFICITY: IN KIDNEY AND IN LIVER FROM JUVENILE AND
CC SEXUALLY MATURE TROUT FROM BOTH SEXES.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U16657; AAA62499.1; -
DR HSSP: P00179; 1D76.
DR InterPro: IPR001128; Cytochrome_P450.
DR Pfam: PF00067; P450.1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450.1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT BINDING 441 441 HEME (BY SIMILARITY).
SQ SEQUENCE 499 AA; 56773 MW; 2D5C106C192ECBA2 CRC64;

Query Match 8.9%; Score 7; DB 1; Length 499;
Best local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 48 LSPTSVP 54
Db 367 LSPTSVP 373

RESULT 7
ZEST_DROME STANDARD; PRT; 574 AA.
ID ZEST_DROME STANDARD; PRT; 574 AA.
AC P09956; Q24596; Q9V3F1;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Regulatory protein zeste.
GN Z OR EG:BACH59J11.3 OR CG7803.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=88142560; PubMed=3125410;
RA Mansukhani A., Gunaratne P.H., Sherwood P.W., Sneath B.J.,
RA Goldberg M.L.;
RT "Nucleotide sequence and structural analysis of the zeste locus of
RT Drosophila melanogaster.";
RL Mol. Gen. Genet. 211:121-128(1988).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN-Oregon-R;
RX MEDLINE=87218538; PubMed=3582372;
RA Pirrotta V., Manet E., Haddon E., Bickel S.E., Benson M.;
RT "Structure and sequence of the Drosophila zeste gene.";
RL EMBO J. 6:791-799(1987).
RN [3]
RN REVISIONS, AND SELF-ASSOCIATION.
RX MEDLINE=90361011; PubMed=2118108;
RA Bickel S.E., Pirrotta V.;
RT "Self-association of the Drosophila zeste protein is responsible for
RT transvection effects.";
RL EMBO J. 9:2959-2967(1990).
RN [4]
RN SEQUENCE FROM N.A.
RP STRAIN-Berkeley;
RC MEDLINE=20196006; PubMed=10731132;
RX

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.M., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Botkova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielson A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Honck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kimms I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q., Ye J.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao O., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [5]
RN SEQUENCE FROM N.A.
RC STRAIN-Oregon-R;
RX MEDLINE=20196011; PubMed=10731137;
RA Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,
RA Barrell B.G., Ferraz C., Vidal S., Brun C., Demallies J., Cadieu E.,
RA Driano S., Gloux S., Lelaune V., Mottier S., Galibert F., Botkova D.,
RA Miñana B., Kafatos F.C., Louis C., Siden-Kimms I., Bolshakov S.,
RA Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablos B.,
RA Modolell J., Peter A., Schoetcler P., Werner M., Moutkioti F.,
RA Beinert N., Dowe G., Schaefer U., Jaecke H., Bucheton A.,
RA Callister D.M., Campbell L.A., Darlamsitou A., Henderson N.S.,
RA McMillan P.J., Sallies C., Tait E.A., Valentini P., Saunders R.D.C.,
RA Glover D.M.;
RT "From sequence to chromosome: the tip of the X chromosome of D.
RT melanogaster.";
RL Science 287:2220-2222(2000).
RN [6]
RN SELF-ASSOCIATION.
RX MEDLINE=93259149; PubMed=8491197;
RA Chen J.D., Pirrotta V.;
RT "Multimerization of the Drosophila zeste protein is required for
RT efficient DNA binding.";
RL EMBO J. 12:2075-2083(1993).
RN [7]
RN SEQUENCE OF 56-323 FROM N.A.
RP STRAIN-ME-K1, ME-K2, ME-L12, ME-NJ1, and ME-NJ2;
RC MEDLINE=93360802; PubMed=8353601;
RX Hey J., Kilman R.M.;
RA "Population genetics and phylogenetics of DNA sequence variation at
RT multiple loci within the Drosophila melanogaster species complex.";
RL Mol. Biol. Evol. 10:804-822(1993).
CC -1- FUNCTION: INVOLVED IN TRANSECTON PHENOMENA (= SYNAPSIS-DEPENDENT
CC GENE EXPRESSION), WHERE THE SYNAPTIC PAIRING OF CHROMOSOMES
CC CARRYING GENES WITH WHICH ZESTE INTERACTS INFLUENCES THE

```
CC      EXPRESSION OF THESE GENES. ZESTIE BINDS TO DNA AND STIMULATES
CC      TRANSCRIPTION FROM A NEARBY PROMOTER.
CC      -1- SUBUNIT: SELF-ASSOCIATE FORMING COMPLEXES OF SEVERAL HUNDRED
CC      MONOMERS.
CC      -1- SUBCELLULAR LOCATION: Nuclear.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
CC      EMBL: Y00049; CAA68262.1; ALT_SEQ.
CC      EMBL: X06743; CAA29918.1; ALT_SEQ.
CC      EMBL: L13043; AAA29026.1; -.
CC      EMBL: L13044; AAA29027.1; -.
CC      EMBL: L13045; AAA29028.1; -.
CC      EMBL: L13046; AAA29029.1; -.
CC      EMBL: L13047; AAA29030.1; -.
CC      EMBL: L13048; AAA29031.1; -.
CC      EMBL: AE003424; AAF45783.1; -.
CC      EMBL: AL133505; CAB63525.1; -.
CC      PIR: A26639; A26639.
CC      PIR: S01272; S01272.
CC      PIR: S12569; S12569.
CC      TRANSFAC: T00918; -.
CC      Flybase: FBgn0004050; Z.
CC      DNA-binding; Transcription regulation; Nuclear protein; Polymorphism.
CC      FT INIT_MER 0 0
CC      FT DOMAIN 1 46 HYDROPHOBIC.
CC      FT DNA_BIND 47 127 SPECIFIC, WITH ZESTIE LOCUS.
CC      FT DOMAIN 5 42 GLY-RICH.
CC      FT DOMAIN 152 430 GLN/ALA-RICH (OPA-REPEAT INVOLVED IN
CC      TRANSCRIPTIONAL ACTIVATION OR REPRESSION
CC      AT DIFFERENT TARGET LOC1) (POTENTIAL).
CC      FT VARIANT 232 233 MISSING (IN STRAIN ME-K2).
CC      FT CONFLICT 496 496 S -> A (IN REF. 4 AND 5).
CC      FT SEQUENCE 574 AA; 61845 MW; 45DCD36CF72F5CF3 CRC64;
CC      -----
CC      Query Match 8.9%; Score 7; DB 1; Length 574;
CC      Best Local Similarity 100.0%; Pred. No. 18;
CC      Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC      QY 9 PAAVPP 15
CC      Db 449 PAAVPP 455
CC      -----
CC      RESULT 8
CC      SL56_HUMAN STANDARD: PRT; 635 AA.
CC      AC Q9Y289;
CC      DT 16-OCT-2001 (Rel. 40; Created)
CC      DT 16-OCT-2001 (Rel. 40; Last sequence update)
CC      DT 16-OCT-2001 (Rel. 40; Last annotation update)
CC      DE Sodium-dependent multivitamin transporter (Na(+)-dependent
CC      DE multivitamin transporter).
CC      GN SLC5A6 OR SMT.
CC      OS Homo sapiens (Human).
CC      OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC      OC Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.
CC      OX NCBI_TaxID=9606;
CC      RN [1]
CC      RP SEQUENCE FROM N.A.
CC      RC TISSUE=Intestine;
CC      RX MEDLINE=99262840; PubMed=10329687;
CC      RA Wang H., Huang W., Fei Y.-J., Xia H., Yang-Feng T.L., Leibach F.H.,
CC      RA Devoe L.D., Ganapathy V., Prasad P.D.;
CC      RA "Human placental Na(+)-dependent multivitamin transporter. Cloning,
CC      RA functional expression, gene structure, and chromosomal localization.";
CC      J. Biol. Chem. 274:14875-14883(1999).
```

```
RN      [2]
RN      SEQUENCE FROM N.A.
RN      RP TISSUE=Intestine;
RN      RC TISSUE=Intestine;
RN      RX MEDLINE=99268779; PubMed=10334869;
RN      RA Prasad P.D., Wang H., Huang W., Fei Y.-J., Leibach F.H., Devoe L.D.,
RN      RA Ganapathy V.;
RN      RA "Molecular and functional characterization of the intestinal Na+-
RN      RA dependent multivitamin transporter.";
RN      RA Arch. Biochem. Biophys. 366:95-106(1999).
CC      CC -1- FUNCTION: TRANSPORTS PANTOTHENATE, BIOTIN AND LIPONATE IN THE
CC      CC PRESENCE OF SODIUM.
CC      CC -1- SIMILARITY: BELONGS TO THE SODIUM:SOLUTE SYMPORTER FAMILY (SSF).
CC      CC -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
CC      EMBL: AF116241; AAD37502.1; -.
CC      EMBL: AF069307; AAD31727.1; -.
CC      EMBL: AF081571; AAD37481.1; -.
CC      GeneW: HGNC:11041; SLC5A6.
CC      DR MIM: 604024; -.
CC      DR InterPro: IPR001734; Na/solut_symport.
CC      DR Pfam: PF00474; SSF: 1.
CC      DR TIGRfams: TIGR00813; sss: 1.
CC      DR PROSITE: PS00456; NA_SOLUT_SYM_1; 1.
CC      DR PROSITE: PS00457; NA_SOLUT_SYM_2; FALSE_NEG.
CC      DR PROSITE: PS00283; NA_SOLUT_SYM_3; 1.
CC      KW Transport; Transmembrane; Sodium transport; Symport; Glycoprotein.
CC      FT TRANSMEM 24 44 POTENTIAL.
CC      FT TRANSMEM 68 88 POTENTIAL.
CC      FT TRANSMEM 101 121 POTENTIAL.
CC      FT TRANSMEM 143 163 POTENTIAL.
CC      FT TRANSMEM 176 196 POTENTIAL.
CC      FT TRANSMEM 199 219 POTENTIAL.
CC      FT TRANSMEM 256 276 POTENTIAL.
CC      FT TRANSMEM 297 317 POTENTIAL.
CC      FT TRANSMEM 336 356 POTENTIAL.
CC      FT TRANSMEM 396 416 POTENTIAL.
CC      FT TRANSMEM 428 448 POTENTIAL.
CC      FT TRANSMEM 456 476 POTENTIAL.
CC      FT TRANSMEM 528 548 POTENTIAL.
CC      FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC      FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC      FT CARBOHYD 498 498 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC      FT SEQUENCE 635 AA; 68701 MW; 4F10369A3916564F CRC64;
CC      -----
CC      Query Match 8.9%; Score 7; DB 1; Length 635;
CC      Best Local Similarity 100.0%; Pred. No. 19;
CC      Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC      QY 46 APLSPS 52
CC      Db 9 APLSPS 15
CC      -----
CC      RESULT 9
CC      GCP3_MOUSE STANDARD: PRT; 677 AA.
CC      AC P58854;
CC      DT 15-JUN-2002 (Rel. 41; Created)
CC      DT 15-JUN-2002 (Rel. 41; Last sequence update)
CC      DT 15-JUN-2002 (Rel. 41; Last annotation update)
CC      DE Gamma-tubulin complex component 3 (GCP-3) (Fragment).
CC      GN GCP3.
CC      OS Mus musculus (Mouse).
CC      OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC      OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Breast tumor;
RA Strausberg R.;
CC Submitted (Mar-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Gamma-tubulin complex is necessary for microtubule
CC nucleation at the centrosome (By similarity).
CC -1- SUBUNIT: Gamma-tubulin complex is composed of gamma-tubulin, GCP2,
CC GCP3, GCP4, GCP5 and GCP6 (By similarity).
CC -1- SUBCELLULAR LOCATION: Centrosome (By similarity).
CC -1- SIMILARITY: BELONGS TO THE GCP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: BC025647; AAB25647.1; -
KW Microtubules.
FT NON_TER 1
FT DOMAIN 595 599 POLY-GLU.
SQ SEQUENCE 677 AA; 78348 MW; 16A85A47357E33C4 CRC64;

Query Match 8.9%; Score 7; DB 1; Length 677;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24 HPVLSFL 30
DB 194 HPVLSFL 200

RESULT 10
CAD2_MOUSE STANDARD; PRT; 906 AA.
ID CAD2_MOUSE STANDARD; PRT; 906 AA.
AC P15116; Q64260;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neural-cadherin precursor (N-cadherin) (Cadherin-2).
GN CDH2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89346748; PubMed=2762814;
RA Miyatani S., Shimamura K., Hatta K., Nagafuchi A., Nose A.,
RA Matsunaga M., Hatta K., Takeichi M.;
RT "Neural cadherin: role in selective cell-cell adhesion.";
RL Science 245:631-635(1989).
RN [2]
RP SEQUENCE FROM N.A.
RA Tamura K.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP PARTIAL SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=92409532; PubMed=1528849;
RA Miyatani S., Copeland N.G., Gilbert D.J., Jenkins N.A., Takeichi M.;
RT "Genomic structure and chromosomal mapping of the mouse N-cadherin
RT gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:8443-8447(1992).
RN [4]
RP DEVELOPMENTAL STAGE.
RC STRAIN=C57BL/6; TISSUE-Testis;
RX MEDLINE=97033837; PubMed=8879495;
RA Munro S.B., Blaschuk O.W.;

```

```

RT "A comprehensive survey of the cadherins expressed in the testes of
RT fetal, immature, and adult mice utilizing the polymerase chain
RT reaction.";
RL Biol. Reprod. 55:822-827(1996).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 160-267.
RX MEDLINE=95191680; PubMed=7885471;
RA Shapiro L., Fannon A.M., Kwong P.D., Thompson A., Lehmann M.S.,
RA Grubel G., Legrand J.-F., Als-Nielsen J., Colman D.R.,
RA Hendrickson W.A.;
RT "Structural basis of cell-cell adhesion by cadherins.";
RL Nature 374:327-337(1995).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (3.4 ANGSTROMS) OF 160-374.
RX MEDLINE=98318235; PubMed=9655503;
RA Tamura K., Shan W.S., Hendrickson W.A., Colman D.R., Shapiro L.;
RT "Structure-function analysis of cell adhesion by neural (N-)
RT cadherin.";
RL Neuron 20:1153-1163(1998).
CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
CC MANNER IN CONNECTING CELLS. CADHERINS MAY THUS CONTRIBUTE TO THE
CC SORTING OF HETEROGENEOUS CELL TYPES. N-CADHERIN MAY BE INVOLVED IN
CC NEURONAL RECOGNITION MECHANISM.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT ALL STAGES OF TESTICULAR
CC DEVELOPMENT WITH HIGHEST LEVELS FOUND IN TESTES OF 21-DAY-OLD
CC MICE.
CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M31131; AAA37353.1; -
DR EMBL: AB008811; BAA23549.1; -
DR EMBL: S45011; AAB23356.1; -
DR PIR: A32759; IJMSCN.
DR PDB: INCG; 10-JUL-95.
DR PDB: INCH; 10-JUL-95.
DR PDB: INCI; 10-JUL-95.
DR PDB: INCI; 18-MAR-99.
DR MGD; MGI:88355; CdH2.
DR InterPro: IPR002126; Cadherin.
DR InterPro: IPR000233; Cadherin_C-term.
DR Pfam: PF00028; cadherin_5.
DR Pfam: PF01049; cadherin_C-term; 1.
DR PRINTS: PR00205; CADHERIN.
DR SMART: SM00112; CA; 5.
DR PROSITE: PS00232; CADHERIN_1; 3.
DR PROSITE: PS50268; CADHERIN_2; 5.
KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
KW Signal; 3D-structure.
FT SIGNAL 1 23
FT PROPEP 24 159
FT CHAIN 160 906 POTENTIAL.
FT DOMAIN 160 724 NEURAL-CADHERIN.
FT TRANSMEM 725 746 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 747 906 POTENTIAL.
FT DOMAIN 160 267 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 268 382 CADHERIN 1.
FT DOMAIN 383 497 CADHERIN 2.
FT DOMAIN 498 603 CADHERIN 3.
FT DOMAIN 604 717 CADHERIN 4.
FT DOMAIN 717 878 CADHERIN 5.
FT DOMAIN 863 878 SER-RICH.
FT CARBOHYD 190 190 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 273 273 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 325 325 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAc. . .) (POTENTIAL).

```

```
FT CARBOHYD 572 572 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 651 651 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 692 692 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 7 9 GRG -> APR (IN REF. 2).
FT CONFLICT 565 565 Y -> N (IN REF. 2).
FT CONFLICT 567 567 Q -> K (IN REF. 2).
FT CONFLICT 624 624 A -> T (IN REF. 2).
SQ SEQUENCE 906 AA; 99761 MW; 0206741B771707B7 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 VAVNLRS 76
Db 117 VAVNLRS 123

RESULT 11
GCP3_HUMAN STANDARD; PRT: 907 AA.
ID GCP3_HUMAN 096CW5: 043631; 060852; 096179;
AC 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Gamma-tubulin complex component 3 (GCP-3) (Spindle pole body protein
GN GCP3
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Cervical carcinoma; PubMed-956967;
RA MEDLINE-98234402; PubMed-956967;
RT Murphy S.M., Urban L., Stearns T.;
RT "The mammalian gamma-tubulin complex contains homologues of the yeast
RL spindle pole body components spc97p and spc98p.";
RL J. Cell Biol. 141:663-674(1998).
[2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND CHARACTERIZATION.
RA MEDLINE-98234404; PubMed-956969;
RT Tassin A.-M., Celati C., Moudjou M., Bornens M.;
RT "Characterization of the human homologue of the yeast spc98p and its
RL association with gamma-tubulin.";
RL J. Cell Biol. 141:689-701(1998).
[3]
RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3).
RC TISSUE=Eye, and Muscle;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Gamma-tubulin complex is necessary for microtubule
CC nucleation at the centrosome.
CC -1- SUBUNIT: Gamma-tubulin complex is composed of gamma-tubulin, GCP2,
CC GCP3, GCP4, GCP5 and GCP6.
CC -1- SUBCELLULAR LOCATION: Centrosome.
CC -1- ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here), 2 and 3; may be
CC produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: Ubiquitously expressed.
CC -1- SIMILARITY: BELONGS TO THE GCP FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
DR EMBL: BC013781; AAH13781.1; -.
DR EMBL: BC007763; AAH07763.1; -.
KW Microtubules; Alternative splicing.
FT DOMAIN 825 829 POLY-GLU
FT VARSPPLIC 817 824 GQMGVTAA -> VEMKLTGV (IN ISOFORM 2).
FT VARSPPLIC 825 907 MISSING (IN ISOFORM 2).
FT VARSPPLIC 391 434 RKGGLASAVHATKTGDPMKSLVQHILSLVSHVPSFLY
FT VARSPPLIC RMT -> PTVRPPTVFPTRDPPTRFDPHVPPTVFPTRV
FT WHISCPTRL (IN ISOFORM 3).
FT VARSPPLIC 435 907 MISSING (IN ISOFORM 3).
FT CONFLICT 208 208 T -> S (IN REF. 2).
FT CONFLICT 361 361 S -> I (IN REF. 2).
SQ SEQUENCE 907 AA; 103570 MW; 70FE2FDB7C80344D CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 HPVLSFL 30
Db 424 HPVLSFL 430

RESULT 12
ST10_HUMAN STANDARD; PRT: 968 AA.
ID ST10_HUMAN 094804: 09UJW4;
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Serine/threonine-protein kinase 10 (EC 2.7.1.37) (Lymphocyte-oriented
GN STK10 OR LOK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RA MEDLINE-99216434; PubMed-10199912;
RA Kuramochi S., Matsuda Y., Okamoto M., Kitamura F., Yonekawa H.,
RA Kuramuyama H.;
RT "Molecular cloning of the human gene STK10 encoding lymphocyte-
RT oriented kinase, and comparative chromosomal mapping of the human,
RL mouse, and rat homologues.";
RL Immunogenetics 49:369-375(1999).
[2]
RP SEQUENCE OF 814-968 FROM N.A.
RC TISSUE=Testis;
RA Bloeker H., Boecher M., Brandt P., Mewes H.-W., Gassenhuber J.,
RA Wiemann S.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CAN ACT ON SUBSTRATES SUCH AS MYELIN BASIC PROTEIN AND
CC HISTONE IIA ON SERINE AND THREONINE RESIDUES (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN LYMPHOID ORGANS.
CC -1- PTM: AUTOPHOSPHORYLATED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC STE20 SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR pfam: PF00069; pkinase; 1.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00220; S_TKC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KM Transferase: Serine/threonine-protein kinase: ATP-binding;
 KW Phosphorylation; Coiled coil.
 FT DOMAIN 36 294 PROTEIN KINASE.
 FT DOMAIN 573 947 COILED COIL (POTENTIAL).
 FT NP_BIND 750 884 GLN-RICH.
 FT BIND 42 50 ATP (BY SIMILARITY).
 FT BIND 65 65 ATP (BY SIMILARITY).
 FT ACT_SITE 157 157 BY SIMILARITY.
 SQ SEQUENCE 968 AA; 112134 MW; 15E245193EC553D CRC64;

Query Match
 Best Local Similarity 100.0%; Score 7; DB 1; Length 968;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 AVVPPLR 17
 Db 405 AVVPPLR 411

RESULT 13
 PSBK_PINTH STANDARD; PRT; 56 AA.

AC P41598;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Photosystem II reaction center protein K precursor (PSII-K).
 GN PSBK.
 OS Pinus thunbergii (Green pine) (Japanese black pine).
 OC Chloroplast.
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
 OX NCBI_TaxID=3350;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95024047; PubMed=7937893;
 RA Waksugi T., Tsudzuki J., Ito S., Nakashima K., Tsudzuki T.,
 RA Sugitara M.;
 RT "Loss of all ndh genes as determined by sequencing the entire
 RT chloroplast genome of the black pine *Pinus thunbergii*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:9794-9798(1994).
 CC -!- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE REACTION CENTER
 CC OF PHOTOSYSTEM II.
 CC -!- SIMILARITY: BELONGS TO THE PSBK FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL: D17510; BAA04312.1; ALT_INIT.
 DR InterPro: IPR003687; PSII_PsbK.
 DR Pfam: PF02533; PsbK; 1.
 KW Photosystem II; Chloroplast.
 FT PROPEP 1 19 POTENTIAL.
 FT CHAIN 20 56 PHOTOSYSTEM II REACTION CENTER PROTEIN K.
 SQ SEQUENCE 56 AA; 6346 MW; 18BIDEF198ACAA5A CRC64;

Query Match
 Best Local Similarity 100.0%; Score 6; DB 1; Length 56;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 25 PVLSFL 30
 Db 39 PVLSFL 44

RESULT 14
 YCXS_OENHO STANDARD; PRT; 111 AA.
 ID YCXS_OENHO

AC Q9MTN3;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical 12.8 kDa protein in ycf9-trns intergenic region (ORF111).
 OS Oenothera hookeri (Hooker's evening primrose).
 OC Chloroplast.
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids II; Myrtales; Onagraceae; Oenothera.
 OX NCBI_TaxID=85636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=cv. Johansen;
 RX MEDLINE=20309318; PubMed=10852478;
 RA Hupfer H., Swiatek M., Hornung S., Herrmann R.G., Maier R.M.,
 RA Chiu W.-T., Sears B.;
 RT "Complete nucleotide sequence of the *Oenothera elata* plastid
 RT chromosome, representing plastome I of the five distinguishable
 RT *Oenothera* plastomes.";
 RL Mol. Gen. Genet. 263:581-585(2000).
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL: AJ271079; CAB67142.1; -
 KW Chloroplast; Hypothetical protein.
 SQ SEQUENCE 111 AA; 12814 MW; E5E0CE989317F140 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 6; DB 1; Length 111;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 VNLSRY 77
 Db 100 VNLSRY 105

RESULT 15
 MF14_MAIZE STANDARD; PRT; 126 AA.
 ID MF14_MAIZE
 AC 001900;

DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE MF514 protein precursor.
 GN MF514.
 OS Zea mays (Maize).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 CC Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=cv. BE10; TISSUE= tassels;
 RX MEDLINE=94004987; PubMed=8401606;
 RX Wright S.Y., Suer M.-M., Bell P.J., Vaundin M., Greenland A.J.;
 RT "Isolation and characterization of male flower cDNAs from maize.";

```

RL Plant J. 3:41-49(1993).
CC -!- TISSUE SPECIFICITY: ENHANCED EXPRESSION IN MALE FLOWERS.
CC ACCUMULATES IN THE TAPETUM.
CC -!- DEVELOPMENTAL STAGE: ASSOCIATED WITH MICROSPOROGENESIS AND
CC DECLINES AS MATURE POLLEN IS PRODUCED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X67323; CAA4737.1; -.
CC PIR: S25104; S25104.
CC MaizeDB: 69183; -.
DR Signal.
KW Signal.
FT SIGNAL 1 23 OR 24, OR 26 (POTENTIAL).
FT CHAIN 24 126 MFS14 PROTEIN.
SQ SEQUENCE 126 AA; 12653 MW; 67F2813AF8C555 CRC64;

```

Query Match 7.6%; Score 6; DB 1; Length 126;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 RPAAP 13
 |||||
 Db 38 RPAAP 43

Search completed: March 4, 2003, 12:45:36
 Job time : 7.39357 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:42:02 ; Search time 8.7249 seconds
(without alignments)
1865.663 Million cell updates/sec

Title: US-09-234-208b-1

Perfect score: 79
Sequence: 1 GTHTSLPPRAAVPVPLRMQP.....VGRGPPDAHVAVNLRYEG 79

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66	83.5	419	09UK79	09UK79 homo sapien
2	8	10.1	217	09MKW6	09MKW6 drosophila
3	8	10.1	481	0945M4	0945M4 arabidopsis
4	8	10.1	520	066518	066518 aquilex aeo
5	8	10.1	865	022774	022774 arabidopsis
6	8	10.1	879	041010	041010 pisum sativ
7	8	10.1	1469	09LKR1	09LKR1 pisum sativ
8	8	10.1	1503	081283	081283 arabidopsis
9	7	8.9	27	09HD19	09HD19 homo sapien
10	7	8.9	39	08WQ02	08WQ02 plasmodium
11	7	8.9	121	08YZW8	08YZW8 anabaena sp
12	7	8.9	146	08UJ38	08UJ38 human immun
13	7	8.9	146	08UJ36	08UJ36 human immun
14	7	8.9	178	09FMC7	09FMC7 arabidopsis
15	7	8.9	200	095PE1	095PE1 amblyomma a
16	7	8.9	241	08TW69	08TW69 methanopyru

17	7	8.9	245	4	014596	014596 homo sapien
18	7	8.9	268	2	09EWB9	09EWB9 streptomyce
19	7	8.9	328	10	081138	081138 arabidopsis
20	7	8.9	350	13	091AL2	091AL2 gallus gall
21	7	8.9	382	5	09UJ31	09UJ31 caenorhabdi
22	7	8.9	389	16	08UEF8	08UEF8 agrobacteri
23	7	8.9	392	12	09Q940	09Q940 shope fibro
24	7	8.9	397	16	0913T2	0913T2 pseudomonas
25	7	8.9	416	16	09KZK5	09KZK5 streptomyce
26	7	8.9	436	16	067814	067814 aquilex aeo
27	7	8.9	445	16	08R814	08R814 thermococci
28	7	8.9	464	16	08YST7	08YST7 anabaena sp
29	7	8.9	469	5	08WQ63	08WQ63 drosophila
30	7	8.9	469	5	08WPA9	08WPA9 drosophila
31	7	8.9	469	5	08WPA5	08WPA5 drosophila
32	7	8.9	471	5	08WQ65	08WQ65 drosophila
33	7	8.9	471	5	08WQ64	08WQ64 drosophila
34	7	8.9	471	5	08WPC6	08WPC6 drosophila
35	7	8.9	471	5	08WPA8	08WPA8 drosophila
36	7	8.9	471	5	08WPA6	08WPA6 drosophila
37	7	8.9	473	5	08WQ66	08WQ66 drosophila
38	7	8.9	473	5	08WPA9	08WPA9 drosophila
39	7	8.9	475	5	08WQ67	08WQ67 drosophila
40	7	8.9	495	2	052799	052799 amycolatops
41	7	8.9	595	5	018900	018900 caenorhabdi
42	7	8.9	635	4	0969Y5	0969Y5 homo sapien
43	7	8.9	798	2	08YR14	08YR14 myxococcus
44	7	8.9	840	16	09X8B0	09X8B0 streptomyce
45	7	8.9	1009	15	0901A3	0901A3 human immun

ALIGNMENTS

RESULT 1	
ID	09UK79
AC	09UK79: PRELIMINARY: PRT: 419 AA.
DT	01-MAY-2000 (TREMBLrel. 13, Created)
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE	Herstatin.
GN	HER-2.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RP	SEQUENCE FROM N.A.
RP	11
RX	MEDLINE-99415951; PubMed-10485918;
RA	Doherty J.K., Bond C., Jardim A., Adelman J.P., Clinton G.M.;
RT	"The HER-2/neu receptor tyrosine kinase gene encodes a secreted
RT	autoinhibitor.";
RT	Proc. Natl. Acad. Sci. U.S.A. 96:10869-10874(1999).
RN	[2]
RP	SEQUENCE FROM N.A.
RP	1
RA	Doherty J.K., Clinton G.M., Adelman J.P., Evans A.J., Hennen W.D.;
RA	Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
DR	EMBL; AF17761; AAD56009.2; -
DR	InterPro: IPR000494; EGFR_L_domain.
DR	InterPro: IPR002174; Furin-like.
DR	Pfam: PF00757; Furin-like; 1.
DR	Pfam: PF01030; Recep_L_domain; 1.
DR	SMART; SM00261; FU; 1.
SQ	SEQUENCE 419 AA: 45472 MW; EC1BE347E2D030C CRC64;
Query Match	
Best local similarity 83.5%; Score 66; DB 4; Length 419;	
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Oy	7 PRPAAVPVPLRMQPGPAHVLISFLRPSWDLVSATYSLPLAPLSPTSPVSVGRGDP 66
Db	347 PRPAAVPVPLRMQPGPAHVLISFLRPSWDLVSATYSLPLAPLSPTSPVSVGRGDP 406

OY 67 DAHAVAV 72
 DB 407 DAHAVAV 412

RESULT 2

O9N6M6 PRELIMINARY: PRT: 217 AA.
 AC O9N6M6:
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE HLHgamma protein.
 GN HLH-GAMMA OR E.
 OS Drosophila hydei (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7224;
 GN [1]
 RP SEQUENCE FROM N.A.
 RA Cooper M.T.D., Tyler D.M., Furrilols M., Chalkiadaki A., Delidakis C.,
 RA Bray S.;
 RT "Spatially restricted factors co-operate with Notch in the regulation
 RT of Enhancer of split genes.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ276313; CAB77019.1; -
 DR Flybase: FBgn012370; Dhyd\HLH-gamma.
 DR InterPro: IPR001092; HLM-basic.
 DR InterPro: IPR003650; Orange.
 DR Pfam: PF00010; HLH; 1.
 DR SMART: SM00353; HLH; 1.
 DR SMART: SM00511; ORANGE; 1.
 DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN.1.
 SQ SEQUENCE 217 AA: 24082 MW: 3917C58F32254ED CRC64;

Query Match 10.1%; Score 8; DB 5; Length 217;
 Best Local Similarity 100.0%; Pred. No. 4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 52 SVPSVPS 59
 DB 159 SVPSVPS 166

RESULT 3

O945M4 PRELIMINARY: PRT: 481 AA.
 AC O945M4:
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE ATG402510/T10P11.19.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 GN [1]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Natuska M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinzaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF412063; AAL06516.1; -
 DR TIGRfams: TIGR00993; 3a0901s04IAP86; 1.

SQ SEQUENCE 481 AA: 53603 MW: D3DD3F8AF6EE7E3A CRC64;

Query Match 10.1%; Score 8; DB 10; Length 481;
 Best Local Similarity 100.0%; Pred. No. 7.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 PAAPVPL 16
 DB 200 PAAPVPL 207

RESULT 4

O66518 PRELIMINARY: PRT: 520 AA.
 AC O66518:
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical protein AO_116.
 GN AO_116.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxID=63563;
 GN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=98196666; PubMed=9537320;
 RX Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjay M., Huber R.,
 RA Felman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 RT aeolicus.";
 RL Nature 392:353-358(1998).
 DR EMBL: AE000674; AAC06482.1; -
 DR InterPro: IPR001450; 4Fe4S_FERREDOXIN.
 DR PROSITE: PS00198; 4Fe4S_FERREDOXIN; UNKNOWN.1.
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 520 AA: 59472 MW: 979DB8B8BF2A20F1 CRC64;

Query Match 10.1%; Score 8; DB 16; Length 520;
 Best Local Similarity 100.0%; Pred. No. 8.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 47 PLSPTSVP 54
 DB 2 PLSPTSVP 9

RESULT 5

O22774 PRELIMINARY: PRT: 865 AA.
 AC O22774:
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Putative chloroplast outer envelope 86-1like protein.
 GN T10P11.19 OR ATG402510.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 GN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=CV. COLOMBIA;
 RC Kaplan N., Johnson D., Schutz K., Gnoj L., Hoffman J., Till S.,
 RA de la Bastide M., Granat S., Hameed A., Gottesman T., Hasegawa A.,
 RA Shohdy N., Parnell L., Dedhia N., Johnson A.F., Lodhi M.,
 RA Martensen R., Chen E.Y., Wilson R., McCombie W.R.;
 RT "Sequence of A. thaliana BAC T10P11 from chromosome IV.";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.

RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,
RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
RA Shekhar M., Schütz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.,
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.,
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC002330; AAC78265.2; -;
DR EMBL: AL161494; CAB80744.1; -;
DR TIGRFAWS: TIGR00993; 3a0901s04IAP86; 1;
SQ SEQUENCE 865 AA; 94388 MW; 5CFD60339AAB39D CRC64;

Query Match 10.1%; Score 8; DB 10; Length 865;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 PAAVPVPL 16
DB 584 PAAVPVPL 591

RESULT 6
O41010 PRELIMINARY; PRT; 879 AA.
AC 041010: 041030;

DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Chloroplast outer envelope protein OEP86 precursor (GTP-binding
DE protein IAP86).
GN IAP86.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;

RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC STRAIN=CV. MIRANDA; TISSUE=LEAF.

RX MEDLINE=95099324; PubMed=7801125;

RA Hirsch S., Muckel E., Heemeyer F., von Heijne G., Soll J.,
RT "A receptor component of the chloroplast protein translocation
RT machinery".

RL Science 266:1989-1992(1994).

RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 551-561; 665-675 AND 867-879.

RX MEDLINE=95063938; PubMed=7973656;

RA Kessler F., Blobel G., Patel H.V., Schnell D.J.,
RT "Identification of two GTP-binding proteins in the chloroplast protein
RT import machinery".

RL Science 266:1035-1039(1994).

CC -1- FUNCTION: INVOLVED IN PROTEIN IMPORT INTO CHLOROPLASTS.

CC -1- SUBCELLULAR LOCATION: CHLOROPLAST OUTER MEMBRANE.

CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-31 IS THE INITIATOR.

DR EMBL: Z31581; CAAB3453.1; -;

DR EMBL: L36857; AAA53276.1; -;

DR TIGRFAWS: TIGR00993; 3a0901s04IAP86; 1.

DR Chloroplast; ATP-binding, outer membrane; GTP-binding.

KT TRANSIT 1 146 CHLOROPLAST.

FT CHAIN 147 879 CHLOROPLAST OUTER ENVELOPE PROTEIN

FT NP_BIND 242 252 ATP (POTENTIAL).

FT CONFLICT 375 GPS -> EOO (IN REF. 2).

SQ SEQUENCE 879 AA; 96548 MW; 3BC1C4E9743A0280 CRC64;

Query Match 10.1%; Score 8; DB 10; Length 879;

Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 PAAVPVPL 16
DB 599 PAAVPVPL 606

RESULT 7
O9LKR1 PRELIMINARY; PRT; 1469 AA.
AC O9LKR1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Chloroplast protein import component Toc159.

OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20177862; PubMed=10712545;
RA Chen K., Chen X., Schnell D.J.,
RT "Initial binding of preproteins involving the toc159 receptor can be
RT bypassed during protein import into chloroplasts.";

RL Plant Physiol. 122:813-822(2000).
DR EMBL: AF262939; AAF75761.1; -;
DR TIGRFAWS: TIGR00993; 3a0901s04IAP86; 1.
SQ SEQUENCE 1469 AA; 158676 MW; AA18A7C3DB5746C4 CRC64;

Query Match 10.1%; Score 8; DB 10; Length 1469;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 PAAVPVPL 16
DB 1189 PAAVPVPL 1196

RESULT 8
O81283 PRELIMINARY; PRT; 1503 AA.
AC O81283;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE T14P8.24 protein.

GN T14P8.24.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Wasnu;
RT "The A. thaliana Genome Sequencing Project.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Kalicki J., Elliott G., Cloud J.,
RT "The sequence of A. thaliana T14P8.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Waterston R.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF069298; AAC19285.1; -;

```

DR  TIGR00993; 3a0901s04IAP86; 1.
SQ  SEQUENCE 1503 AA; 160818 MW; 04AEDE84C1BED3f6 CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 10; Length 1503;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  9 PAAVPL 16
    |||||
Db  1222 PAAVPL 1229

RESULT 9
O9HD19 PRELIMINARY; PRT; 27 AA.
AC  O9HD19;
DT  01-MAR-2001 (TREMBLrel. 16, Created)
DT  01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE  01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE  Sodium-dependent multivitamin transporter (Fragment).
GN  SMVT OR SLC5A6.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=HEART;
RA  Rubin S.A.; Dey S.; Chatterjee N.; Said H.M.;
RT  "Molecular characterization of the human heat SMVT cDNA.";
RN  Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RX  MEDLINE-21952373; PubMed-11955628;
RX  Dey S.; Subramanian V.S.; Chatterjee N.S.; Rubin S.A.; Said H.M.;
RT  "Characterization of the 5' regulatory region of the human sodium-
RT  dependent multivitamin transporter, hSMVT.";
RL  Biochim. Biophys. Acta 1574:187-192(2002).
DR  EMBL; AF288781; AAG00587.1; -
DR  EMBL; AF442149; AAL84706.1; -
DR  EMBL; AF442150; AAL84707.1; -
FT  NON_TER 27
SQ  SEQUENCE 27 AA; 2620 MW; 27CDDC338BD55E9D CRC64;

Query Match
Best Local Similarity 8.9%; Score 7; DB 4; Length 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  46 APLSPTS 52
    |||||
Db  9 APLSPTS 15

RESULT 10
O8WQJ2 PRELIMINARY; PRT; 39 AA.
AC  O8WQJ2;
DT  01-MAR-2002 (TREMBLrel. 20, Created)
DT  01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE  01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE  RhoP-3 (Fragment).
OS  Plasmodium falciparum.
OC  Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX  NCBI_TaxID=5833;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=PF452;
RA  Wang T.; Sam-Yellowe T.Y.;
RT  "Conservation of the rhopty protein RhoP-3 among human and rodent
RT  Plasmodium species and its potential as a malaria vaccine candidate.";
RL  Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AY044910; AAL65267.1; -
FT  NON_TER 1

```

```

FT  NON_TER 39
SQ  SEQUENCE 39 AA; 4498 MW; 853A12A3950B87A7 CRC64;

Query Match
Best Local Similarity 8.9%; Score 7; DB 5; Length 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  54 PISPSV 60
    |||||
Db  15 PISPSV 21

RESULT 11
O8YZW8 PRELIMINARY; PRT; 121 AA.
AC  O8YZW8;
DT  01-MAR-2002 (TREMBLrel. 20, Created)
DT  01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE  01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE  Hypothetical protein A110337.
GN  A110337.
OS  Anabaena sp. (strain PCC 7120).
OC  Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX  NCBI_TaxID=103690;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE-21595285; PubMed-11759840;
RA  Kaneko T.; Nakamura Y.; Molk C.P.; Kuritz T.; Sasamoto S.;
RA  Watanabe A.; Itiguchi M.; Ishikawa A.; Kawashima K.; Kimura T.;
RA  Kishida Y.; Kohara M.; Matsumoto M.; Matsuno A.; Muraki A.;
RA  Nakazaki N.; Shimpo S.; Sugimoto M.; Takazawa M.; Yamada M.;
RA  Yasuda M.; Tabata S.;
RT  "Complete genomic sequence of the filamentous nitrogen-fixing
RT  cyanobacterium Anabaena sp. strain PCC 7120.";
RL  DNA Res. 8:205-213(2001).
DR  EMBL; AP003582; BAB72293.1; -
DR  InterPro: IPR003477; Psmk.
DR  Pfam: PF02452; Psmk; 1.
KW  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 121 AA; 13266 MW; 8E01290F333B1B7A CRC64;

Query Match
Best Local Similarity 8.9%; Score 7; DB 16; Length 121;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  25 PVLSPFLR 31
    |||||
Db  56 PVLSPFLR 62

RESULT 12
O8UJ38 PRELIMINARY; PRT; 146 AA.
AC  O8UJ38;
DT  01-MAR-2002 (TREMBLrel. 20, Created)
DT  01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE  01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE  POL protein (Fragment).
GN  POL.
OS  Human immunodeficiency virus type 1.
OC  Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX  NCBI_TaxID=11676;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=98CMA307;
RX  MEDLINE-21565561; PubMed-11709100;
RA  Luk K.-C.; Kaplue L.; Zekeng L.; Soriano V.; Gurtler L.; Devare S.G.;
RA  Schochetman G.; Hackett J. Jr.;
RT  "Naturally Occurring Sequence Polymorphisms within HIV Type 1 Group O
RT  Protease.";
RL  AIDS Res. Hum. Retroviruses 17:1555-1561(2001).
DR  EMBL; AF380212; AAL32185.1; -
DR  InterPro: IPR001995; Aspprotease_ftrv.

```

```

DR InterPro: IPR001969; Aspprotease_site.
DR Pfam: PF00077; rvp; 1.
DR PROSITE: PS00141; ASP_PROTEASE; UNKNOWN_1.
DR PROSITE: PS50175; ASP_PROT_RETROV; 1.
FT NON_TER 1
FT CHAIN 40 >138 PROTEASE.
FT NON_TER 146 146
SQ SEQUENCE 146 AA; 15018 MW; 70C5BA6539C7534 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 TSVPISP 57
   |||||
Db 7 TSVPISP 13

RESULT 13
O8U36 PRELIMINARY; PRT; 146 AA.
ID O8U36:
AC O8U36:
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Pol protein (Fragment).
GN POL.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxId=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98CMA376;
RX MEDLINE=21565561; PubMed=11709100;
RA Luk K.-C., Kapteu L., Zekeng L., Soriano V., Gurtler L., Devare S.G.,
RA Schocherman G., Hackelt J. Jr.;
RT "Naturally Occurring Sequence Polymorphisms within HIV Type 1 Group O
RT Protease.";
RL AIDS Res. Hum. Retroviruses 17:1555-1561(2001).
DR EMBL: AF380214; AAL32187.1; -.
DR InterPro: IPR001995; Aspprotease_rtrv.
DR InterPro: IPR001969; Aspprotease_site.
DR Pfam: PF00077; rvp; 1.
DR PROSITE: PS00141; ASP_PROTEASE; UNKNOWN_1.
DR PROSITE: PS50175; ASP_PROT_RETROV; 1.
FT NON_TER 1
FT CHAIN 40 >138 PROTEASE.
FT NON_TER 146 146
SQ SEQUENCE 146 AA; 15127 MW; F498137384AB00D6 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 TSVPISP 57
   |||||
Db 7 TSVPISP 13

RESULT 14
O9FMC7 PRELIMINARY; PRT; 178 AA.
ID O9FMC7:
AC O9FMC7:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 20.3 kDa protein.
OS Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxId=3702;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RX MEDLINE=98162728; PubMed=9501997;
RA Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. III.
RT Sequence features of the regions of 1,191,918 bp covered by seventeen
RT physically assigned pl clones.";
RL DNA Res. 4:401-414(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carinci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Full length cDNA of gene MUK1.15/AT5g04830 (GI:9758458).";
RT Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Pham P.K., Banh J., Banno F., Dale J.M.,
RA Goldsmith A.D., Jiang P.X., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Yamamura Y., Yu G., Yu S., Bowser L.,
RA Carinci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E., Lam B.,
RA Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai K., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "Full length cDNA of gene MUK1.15/AT5g04830 (GI:9758458).";
RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB008271; BAB08987.1; -.
DR EMBL: AF360335; AAK28632.1; -.
DR EMBL: AY051070; AAK93747.1; -.
KW Hypothetical protein.
SQ SEQUENCE 178 AA; 20294 MW; 955AC984A407FF54 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 SAFYSLP 44
   |||||
Db 66 SAFYSLP 72

RESULT 15
O95PE1 PRELIMINARY; PRT; 200 AA.
ID O95PE1:
AC O95PE1:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 21.7 kDa protein (Fragment).
OS Amblyomma americanum.
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Amblyomma.
OX NCBI_TaxId=6943;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MIDGUT.
RA Jaworski D.C., Barbour A.G.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF133719; AAK49812.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 200 200
SQ SEQUENCE 200 AA; 21688 MW; 6ABD8ADFCA3B84 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 4 SLPRPA 10
| | | | |
Db 124 SLPRPA 130

Search completed: March 4, 2003, 12:46:38
Job time : 10.7249 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:43:52 ; Search time 6.34538 Seconds
(without alignments)
366.315 Million cell updates/sec

Title: US-09-234-208B-1

Sequence: 1 GTHSLPRPAAPVPLRMQP.....VGRGPDPAHVAVNLRYEG 79

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents-AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PTCUS.COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	79	US-09-630-155-1	Sequence 1, Appli
2	79	100.0	419	US-09-630-155-2	Sequence 2, Appli
3	7	8.9	328	US-09-300-672-2	Sequence 2, Appli
4	7	8.9	635	US-09-014-969-11	Sequence 11, Appli
5	7	8.9	1091	US-09-306-595C-7	Sequence 7, Appli
6	6	7.6	12	US-08-811-492-146	Sequence 146, App
7	6	7.6	15	US-08-036-555B-49	Sequence 49, Appl
8	6	7.6	15	US-08-469-569-49	Sequence 49, Appl
9	6	7.6	15	US-08-249-322A-49	Sequence 49, Appl
10	6	7.6	15	US-08-469-526A-49	Sequence 49, Appl
11	6	7.6	15	US-08-734-591A-49	Sequence 49, Appl
12	6	7.6	15	US-08-469-660-49	Sequence 49, Appl
13	6	7.6	15	US-08-341-018-82	Sequence 82, Appl
14	6	7.6	15	US-08-470-335-49	Sequence 49, Appl
15	6	7.6	15	US-08-735-021-49	Sequence 49, Appl
16	6	7.6	15	US-08-734-664A-49	Sequence 49, Appl
17	6	7.6	15	US-08-470-339-49	Sequence 49, Appl
18	6	7.6	15	US-08-467-602-49	Sequence 49, Appl
19	6	7.6	15	PCT-US94-05083C-46	Sequence 46, Appl
20	6	7.6	15	PCT-US95-06846A-49	Sequence 49, Appl
21	6	7.6	16	US-08-036-555B-37	Sequence 37, Appl
22	6	7.6	16	US-08-469-569-37	Sequence 37, Appl
23	6	7.6	16	US-08-249-322A-37	Sequence 37, Appl
24	6	7.6	16	US-08-469-526A-37	Sequence 37, Appl
25	6	7.6	16	US-08-734-591A-37	Sequence 37, Appl
26	6	7.6	16	US-08-469-660-37	Sequence 37, Appl
27	6	7.6	16	US-08-470-335-37	Sequence 37, Appl

28	6	7.6	16	US-08-735-021-37	Sequence 37, Appl
29	6	7.6	16	US-08-734-664A-37	Sequence 37, Appl
30	6	7.6	16	US-08-470-339-37	Sequence 37, Appl
31	6	7.6	16	US-08-467-602-37	Sequence 37, Appl
32	6	7.6	16	PCT-US94-05083C-37	Sequence 37, Appl
33	6	7.6	16	PCT-US95-06846A-37	Sequence 37, Appl
34	6	7.6	34	US-08-118-270-81	Sequence 81, Appl
35	6	7.6	34	US-08-085-127-11	Sequence 11, Appl
36	6	7.6	34	US-08-319-052-21	Sequence 21, Appl
37	6	7.6	34	US-08-442-108B-21	Sequence 21, Appl
38	6	7.6	34	PCT-US93-08528-81	Sequence 81, Appl
39	6	7.6	74	US-09-134-001C-4092	Sequence 4092, Ap
40	6	7.6	145	US-08-565-386-15	Sequence 15, Appl
41	6	7.6	166	US-09-615-192A-273	Sequence 273, App
42	6	7.6	290	US-09-386-653A-7	Sequence 7, Appl
43	6	7.6	295	US-08-411-705-4	Sequence 4, Appl
44	6	7.6	304	US-09-088-651-2	Sequence 2, Appl
45	6	7.6	310	US-09-199-637A-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-09-630-155-1
; Sequence 1, Application US/09630155
; Patent No. 6414130
; GENERAL INFORMATION:
; APPLICANT: Doherty, Joni Kristin and Gail M. Clinton
; TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVIS WRIGHT TREMAINE LLP
; STREET: 1501 Fourth Avenue, 2600 Century Square
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: PC compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/630,155
; FILING DATE: 16-Jan-2001
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Davison, Barry L.
; REGISTRATION NUMBER: 47,309
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206 628-7621
; TELEFAX: 206 628-7699
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: HER-2 ECD antagonist
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-630-155-1
Query Match 100.0%; Score 79; DB 4; Length 79;
Best Local Similarity 100.0%; Pred. No. 2.2e-69;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTHSLPRPAAPVPLRMQPGAHVPLSLRPSMDLSAFYSLPLAPLSPTVPIPSVSV 60
DB 1 GTHSLPRPAAPVPLRMQPGAHVPLSLRPSMDLSAFYSLPLAPLSPTVPIPSVSV 60
QY 61 GRGPPDAHVAHVNLRYEG 79

DB 61 GRGPPDAHVAVNLSRYEG 79

RESULT 2

US-09-630-155-2
Sequence 2, Application US/09630155
Patent No. 6414130

GENERAL INFORMATION:

APPLICANT: Doherty, Joni Kristin and Gail M. Clinton
TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: DAVIS WRIGHT TREMAINE LLP
STREET: 1501 Fourth Avenue, 2600 Century Square
CITY: Seattle
STATE: Washington
COUNTRY: U.S.A.
ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: PC compatible
OPERATING SYSTEM: Windows95
SOFTWARE: Word

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/630,155
FILING DATE: 16-Jan-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Davison, Barry L.
REGISTRATION NUMBER: 47,309
REFERENCE/DOCKET NUMBER: 49321-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206 628-7621
TELEFAX: 206 628-7699

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 419
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown

MOLECULE TYPE: polypeptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-630-155-2

Query Match 100.0%; Score 79; DB 4; Length 419;
Best Local Similarity 100.0%; Pred. No. 9.8e-69;

Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTHTLPRPAVVPPLRMQGPAPVLSFLRPSMDLVSAFYSLPLAPLSPTSPVSPSV 60

DB 341 GTHTLPRPAVVPPLRMQGPAPVLSFLRPSMDLVSAFYSLPLAPLSPTSPVSPSV 400

QY 61 GRGPPDAHVAVNLSRYEG 79

DB 401 GRGPPDAHVAVNLSRYEG 419

RESULT 3

US-09-300-672-2
Sequence 2, Application US/09300672
Patent No. 6248937

GENERAL INFORMATION:

APPLICANT: Finkelstein, Ruth R.
APPLICANT: Lynch, Tim
APPLICANT: Goodman, Howard M.
APPLICANT: Wang, Ming-Li

TITLE OF INVENTION: A TRANSCRIPTION FACTOR REGULATING SEED DEVELOPMENT,
FILE REFERENCE: 480.89(HV)
CURRENT APPLICATION NUMBER: US/09/300,672

CURRENT FILING DATE: 1999-04-27
NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 2
LENGTH: 328
TYPE: PRT
ORGANISM: Arabidopsis
US-09-300-672-2

Query Match 8.9%; Score 7; DB 4; Length 328;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LLPRPA 11

DB 144 LLPRPA 150

RESULT 4

US-09-014-969-11
Sequence 11, Application US/09014969
Patent No. 5965397

GENERAL INFORMATION:

APPLICANT: McCoy, John M.
APPLICANT: Lavallee, Edward R.
APPLICANT: Racie, Lisa A.
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
APPLICANT: Agostino, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
NUMBER OF SEQUENCES: 32
ENCODING THEM
CORRESPONDENCE ADDRESS:
ADDRESS: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/014,969
FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:
LENGTH: 635 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear

MOLECULE TYPE: protein
US-09-014-969-11

Query Match 8.9%; Score 7; DB 2; Length 635;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 APLSPS 52

DB 9 APLSPS 15

RESULT 5

US-09-306-595C-7
; Sequence 7, Application US/09306595C
; Patent No. 6284506
; GENERAL INFORMATION:
; APPLICANT: HOSHINO, Tatsuo
; APPLICANT: OIIMA, Kazuyuki
; APPLICANT: SETOGUCHI, Yutaka
; TITLE OF INVENTION: ISOPRENOID PRODUCTION
; FILE REFERENCE: ISOPRENOID PRODUCTION
; CURRENT APPLICATION NUMBER: US/09/306,595C
; CURRENT FILING DATE: 1999-05-06
; PRIOR APPLICATION NUMBER: 98108210
; PRIOR FILING DATE: 1998-05-06
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 1091
; TYPE: PRT
; ORGANISM: Phafila rhodozyma
US-09-306-595C-7

Query Match 8.9%; Score 7; DB 4; Length 1091;
Best local Similarity 100.0%; Pred. No. 83;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 PISPSV 60
|111111|
DB 506 PISPSV 512

RESULT 6
US-08-811-492-146
; Sequence 146, Application US/08811492
; Patent No. 5834247

GENERAL INFORMATION:

APPLICANT: COMB, DONALD G.
APPLICANT: PERLER, FRANCINE B.
APPLICANT: JACK, WILLIAM E.
APPLICANT: XU, MING-OUN
APPLICANT: HODGES, ROBERT A.
APPLICANT: NOREN, CHRISTOPHER J.
APPLICANT: CHONG, SHAORONG S.C.
APPLICANT: ADAM, ERIC
APPLICANT: SOUTHWORTH, MAURICE
TITLE OF INVENTION: MODIFIED PROTEINS, METHODS OF THEIR
TITLE OF INVENTION: PRODUCTION AND METHODS FOR PURIFICATION OF TARGET
TITLE OF INVENTION: PROTEINS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: GREGORY D. WILLIAMS; NEW ENGLAND BIOLABS, INC.
STREET: 32 TOZER ROAD
CITY: BEVERLY
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 01915

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC\DOS\MS\DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,492
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/580,555
FILING DATE: 29-DEC-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/496,247
FILING DATE: 28-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/146,885
FILING DATE: 03-NOV-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/004,139
FILING DATE: 09-DEC-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Gregory D
REGISTRATION NUMBER: 30901
REFERENCE/DOCKET NUMBER: NEB-036C4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 508-927-5054
TELEFAX: 509-927-1705
TELEX:
INFORMATION FOR SEQ ID NO: 146:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-811-492-146

Query Match 7.6%; Score 6; DB 2; Length 12;
Best local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 LAPLSP 50
|11111|
DB 7 LAPLSP 12

RESULT 7
US-08-036-555B-49
; Sequence 49, Application US/08036555B
; Patent No. 5530109

GENERAL INFORMATION:

APPLICANT: Goodearl, Andrew; Stroobant, Paul;
APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;
APPLICANT: Chen, Maio Su; Hiles, Ian
TITLE OF INVENTION: Glial Mitogenic Factors, Their
TITLE OF INVENTION: Preparation and Use
NUMBER OF SEQUENCES: 184
CORRESPONDENCE ADDRESS:
ADDRESSEE: Feltz & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/036,555B
FILING DATE: 24-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/965,173
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/940,389
FILING DATE: 03-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/907,138
FILING DATE: 30-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/863,703
FILING DATE: 03-APRIL-1992
PRIOR APPLICATION DATA:

APPLICATION NUMBER: U.K. 91 07566.3
FILING DATE: 10-APRIL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Tsai, Christine H.
REGISTRATION NUMBER: 34,266
REFERENCE/DOCKET NUMBER: LUD 5250.4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 15
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-036-555B-49

Query Match 7.6%; Score 6; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 SPVSVG 61
|||||
DB 2 SPVSVG 7

RESULT 8
US-08-469-569-49
Sequence 49, Application US/08469569
Patent No. 5606032
GENERAL INFORMATION:
APPLICANT: Goodearl, Andrew; Stroobant, Paul;
APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionl, Mark;
APPLICANT: Chen, Malo Su; Hiles, Ian
TITLE OF INVENTION: Glial Mitogenic Factors, Their
NUMBER OF INVENTION: Preparation and Use
NUMBER OF SEQUENCES: 184
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,569
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/036,555
FILING DATE: 24-MAR-1993
APPLICATION NUMBER: 07/965,173
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/940,389
FILING DATE: 03-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/907,138
FILING DATE: 30-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/863,703
FILING DATE: 03-APRIL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.K. 91 07566.3
FILING DATE: 10-APRIL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Tsai, Christine H.
REGISTRATION NUMBER: 34,266

REFERENCE/DOCKET NUMBER: LUD 5250.4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 15
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-469-569-49

Query Match 7.6%; Score 6; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 SPVSVG 61
|||||
DB 2 SPVSVG 7

RESULT 9
US-08-249-322A-49
Sequence 49, Application US/08249322A
Patent No. 5716930
GENERAL INFORMATION:
APPLICANT: Goodearl, Andrew; Stroobant, Paul;
APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionl, Mark;
APPLICANT: Chen, Malo Su; Hiles, Ian
TITLE OF INVENTION: Glial Mitogenic Factors, Their
NUMBER OF INVENTION: Preparation and Use
NUMBER OF SEQUENCES: 184
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/249,322A
FILING DATE: 26-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/036,555
FILING DATE: 24-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/965,173
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/940,389
FILING DATE: 03-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/907,138
FILING DATE: 30-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/863,703
FILING DATE: 03-APRIL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.K. 91 07566.3
FILING DATE: 10-APRIL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Tsai, Christine H.
REGISTRATION NUMBER: 34,266
REFERENCE/DOCKET NUMBER: LUD 250.4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884

;; INFORMATION FOR SEQ ID NO: 49:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 15
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
US-08-249-322A-49

Query Match 7.6%; Score 6; DB 1;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 SPVSVG 61
|||||
Db 2 SPVSVG 7

RESULT 10
US-08-469-526A-49
; Sequence 49, Application US/08469526A
; Patent No. 5792849
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew
; APPLICANT: Stroobant, Paul
; APPLICANT: Minghetti, Luisa
; APPLICANT: Waterfield, Michael
; APPLICANT: Marchionni, Mark
; APPLICANT: Chen, Mario Su
; APPLICANT: Hiles, Ian
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
; TITLE OF INVENTION: PREPARATION AND USE
; NUMBER OF SEQUENCES: 187
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,526A
; FILING DATE: 06 June 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/036,555
; FILING DATE: 24-MAR-1993
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 03-JUN-1992
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APRIL-1992
; APPLICATION NUMBER: U.K. 91 07566.3
; FILING DATE: 10-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Bieker-Brady, Kristina
; REGISTRATION NUMBER: 39,109
; REFERENCE/DOCKET NUMBER: 04585/00200A
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15
; TYPE: amino acid
; STRANDEDNESS: single

;; TOPOLOGY: linear
US-08-469-526A-49

Query Match 7.6%; Score 6; DB 1;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 SPVSVG 61
|||||
Db 2 SPVSVG 7

RESULT 11
US-08-734-591A-49
; Sequence 49, Application US/08734591A
; Patent No. 5854220
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew
; APPLICANT: Stroobant, Paul
; APPLICANT: Minghetti, Luisa
; APPLICANT: Waterfield, Michael
; APPLICANT: Marchionni, Mark
; APPLICANT: Hiles, Ian
; APPLICANT: Chen, Mario
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
; TITLE OF INVENTION: PREPARATION AND USE
; NUMBER OF SEQUENCES: 187
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM Compatible Pentium
; OPERATING SYSTEM: Windows95
; SOFTWARE: Wordperfect (Version 7.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/734,591A
; FILING DATE: 22-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/470,335
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/036,555
; FILING DATE: 03-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 30-JUN-1992
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APR-1992
; APPLICATION NUMBER: UK 91 07566.3
; FILING DATE: 10-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Bieker-Brady, Kristina
; REGISTRATION NUMBER: 39,109
; REFERENCE/DOCKET NUMBER: 04585/00200P
; TELEPHONE: (617) 428-0200
; TELEFAX: (617) 428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 49:

SEQUENCE CHARACTERISTICS:
LENGTH: 15
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-734-591A-49

Query Match 7.6%; Score 6; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 SPVSVG 61
|||||
DB 2 SPVSVG 7

RESULT 12
US-08-469-660-49
Sequence 49, Application US/08469660
Patent No. 5876973

GENERAL INFORMATION:
APPLICANT: Gwynne, David I.; Marchionni, Mark;
APPLICANT: McBurney, Robert N.
TITLE OF INVENTION: INHIBITORS OF CELL PROLIFERATION,
TITLE OF INVENTION: THEIR PREPARATION AND USE
NUMBER OF SEQUENCES: 184
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
ZIP: 0211-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,660
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/011,396
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/984,085
FILING DATE: 01-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/951,747
FILING DATE: 25-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/927,337
FILING DATE: 10-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 04585/017004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: 200154

INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 15
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-469-660-49

Query Match 7.6%; Score 6; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 SPVSVG 61

DB 2 SPVSVG 7
|||||

RESULT 13
US-08-341-018-82
Sequence 82, Application US/08341018A
Patent No. 6087323

GENERAL INFORMATION:
APPLICANT: Gwynne, David I.
APPLICANT: Mahanthappa, Nagesh K.
APPLICANT: Marchionni, Mark A.
APPLICANT: Birmingham-McDonogh, Olivia
APPLICANT: Goldney, Stanley M.
APPLICANT: McBurney, Robert N.

TITLE OF INVENTION: USE OF NEUREGULINS AS MODULATORS OF
TITLE OF INVENTION: CELLULAR COMMUNICATION
FILE REFERENCE: 04585/041001
CURRENT APPLICATION NUMBER: US/08/341,018A
CURRENT FILING DATE: 1994-11-17
NUMBER OF SEQ ID NOS: 87
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 82

LENGTH: 15
TYPE: PRP
ORGANISM: Bos taurus
US-08-341-018-82

Query Match 7.6%; Score 6; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 SPVSVG 61
|||||
DB 2 SPVSVG 7

RESULT 14
US-08-470-335-49

Sequence 49, Application US/08470335F
Patent No. 6147190
GENERAL INFORMATION:
APPLICANT: GOODEARL, ANDREW
APPLICANT: STROOBANT, PAUL
APPLICANT: MINGHETTI, LUISA
APPLICANT: WATERFIELD, MICHAEL
APPLICANT: MARCHIONNI, MARK
APPLICANT: CHEN, MARIO S.

TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
TITLE OF INVENTION: PREPARATION AND USE
FILE REFERENCE: 04585/00200B
CURRENT APPLICATION NUMBER: US/08/470,335F
CURRENT FILING DATE: 1995-06-06
EARLIER APPLICATION NUMBER: 08/036,555
EARLIER FILING DATE: 1993-03-24
NUMBER OF SEQ ID NOS: 252
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 49

LENGTH: 15
TYPE: PRP
ORGANISM: Bos taurus
US-08-470-335-49

Query Match 7.6%; Score 6; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 SPVSVG 61
|||||
DB 2 SPVSVG 7

```
RESULT 15
US-08-735-021-49
; Sequence 49, Application US/08735021B
; Patent No. 6194377
; GENERAL INFORMATION:
; APPLICANT: GOODEARL, ANDREW
; APPLICANT: STROBANT, PAUL
; APPLICANT: MINGHETTI, LUISA
; APPLICANT: WATERFIELD, MICHAEL
; APPLICANT: MARCHIONNI, MARK
; APPLICANT: CHEN, MARIO S.
; APPLICANT: HILES, IAN
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
; FILE REFERENCE: 04585/00200L
; CURRENT APPLICATION NUMBER: US/08/735,021B
; EARLIER APPLICATION NUMBER: 08/472,065
; EARLIER FILING DATE: 1995-06-06
; EARLIER APPLICATION NUMBER: 08/036,555
; EARLIER FILING DATE: 1993-03-24
; EARLIER APPLICATION NUMBER: 07/965,173
; EARLIER FILING DATE: 1992-10-23
; EARLIER APPLICATION NUMBER: 07/940,389
; EARLIER FILING DATE: 1992-09-03
; EARLIER APPLICATION NUMBER: 07/907,138
; EARLIER FILING DATE: 1992-06-30
; EARLIER APPLICATION NUMBER: 07/863,703
; EARLIER FILING DATE: 1992-04-03
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 49
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Bos taurus
US-08-735-021-49

Query Match          7.6%; Score 6; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 SPVSVG 61
   |||||
Db  2 SPVSVG 7

Search completed: March 4, 2003, 12:48:27
Job time : 7.34538 secs
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:46:47 ; Search time 3.3133 Seconds
(without alignments)
1000.035 Million cell updates/sec

Title: US-09-234-208B-1

Perfect score: 79
Sequence: 1 GTHSLPRPAAVPVPLRMQP.....VGRCPDPAHVAVNLSRYEG 79

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 188354 seqs, 42170167 residues

Word size : 0

Total number of hits satisfying chosen parameters: 188354

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/PCT05_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	8.9	61	10 US-09-864-761-44740	Sequence 44740, A
2	7	8.9	215	9 US-10-001-876-210	Sequence 210, App
3	7	8.9	824	9 US-09-909-567B-53	Sequence 53, Appl
4	7	8.9	880	9 US-09-893-519A-36	Sequence 36, Appl
5	7	8.9	906	10 US-09-746-491-48	Sequence 48, Appl
6	6	7.6	30	10 US-09-864-761-45510	Sequence 45510, A
7	6	7.6	34	9 US-09-755-109-21	Sequence 21, Appl
8	6	7.6	38	10 US-09-864-761-41953	Sequence 41953, A
9	6	7.6	68	10 US-09-764-887-226	Sequence 226, App
10	6	7.6	77	10 US-09-864-761-35919	Sequence 35919, A
11	6	7.6	86	10 US-09-764-869-1118	Sequence 1118, App
12	6	7.6	93	10 US-09-764-869-798	Sequence 798, App
13	6	7.6	114	10 US-09-864-761-38224	Sequence 38224, A
14	6	7.6	131	10 US-09-771-161A-148	Sequence 148, App
15	6	7.6	147	10 US-09-764-847-828	Sequence 828, App
16	6	7.6	168	9 US-09-738-626-5722	Sequence 5722, App
17	6	7.6	168	10 US-09-925-301-1470	Sequence 1470, App
18	6	7.6	238	12 US-10-024-579-16	Sequence 16, Appl
19	6	7.6	243	9 US-09-989-442-122	Sequence 122, Appl

20	6	7.6	257	12 US-10-024-579-14	Sequence 14, Appl
21	6	7.6	264	10 US-09-848-696-4	Sequence 4, Appl
22	6	7.6	264	12 US-10-024-579-12	Sequence 12, Appl
23	6	7.6	267	9 US-09-808-602-23	Sequence 23, Appl
24	6	7.6	267	9 US-09-808-602-25	Sequence 25, Appl
25	6	7.6	283	12 US-10-024-579-10	Sequence 10, Appl
26	6	7.6	290	9 US-10-041-006A-7	Sequence 7, Appl
27	6	7.6	290	9 US-10-028-072-222	Sequence 222, App
28	6	7.6	290	9 US-10-121-049-222	Sequence 222, App
29	6	7.6	290	9 US-10-123-904-222	Sequence 222, App
30	6	7.6	290	9 US-10-140-470-222	Sequence 222, App
31	6	7.6	290	9 US-10-175-746-222	Sequence 222, App
32	6	7.6	290	9 US-10-176-918-222	Sequence 222, App
33	6	7.6	290	9 US-10-176-921-222	Sequence 222, App
34	6	7.6	290	9 US-10-137-865-222	Sequence 222, App
35	6	7.6	290	9 US-10-140-474-222	Sequence 222, App
36	6	7.6	290	9 US-10-142-431-222	Sequence 222, App
37	6	7.6	290	9 US-10-143-114-222	Sequence 222, App
38	6	7.6	290	9 US-10-140-002-222	Sequence 222, App
39	6	7.6	290	12 US-10-040-655-7	Sequence 7, Appl
40	6	7.6	298	10 US-09-764-853-552	Sequence 552, App
41	6	7.6	299	9 US-09-808-602-63	Sequence 63, Appl
42	6	7.6	306	12 US-10-080-644-11	Sequence 11, Appl
43	6	7.6	310	9 US-09-975-719-11	Sequence 49, Appl
44	6	7.6	316	10 US-09-799-777-49	Sequence 160, App
45	6	7.6	323	9 US-09-984-245-160	

ALIGNMENTS

RESULT 1

US-09-864-761-44740

Sequence 44740, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: Aemolca-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263,6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 44740
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC016057.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.72
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.59
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.53
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.52
; OTHER INFORMATION: SWISSPROT HIT: Q13563, EVALUATE 2.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BF570694.1, EVALUATE 4.00e-16
US-09-864-761-44740
```

```

Query Match      8.9%; Score 7; DB 10; Length 61;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 23 AHPVLSF 29
DB 37 AHPVLSF 43
```

```

RESULT 2
US-10-001-876-210
; Sequence 210, Application US/10001876
; Patent No. US20020177140A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Hervé
; APPLICANT: Caferkey, Robert
; APPLICANT: Ali, Shujath
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenchua
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
; FILE REFERENCE: DEX-0285
; CURRENT APPLICATION NUMBER: US/10/001,876
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,186
; PRIOR FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 210
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-876-210
```

```

Query Match      8.9%; Score 7; DB 9; Length 215;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 27 LSFRLPS 33
DB 60 LSFRLPS 66
```

```

RESULT 3
US-09-909-567B-53
; Sequence 53, Application US/09909567B
; Publication No. US20030022257A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Macina, Roberto A.
; APPLICANT: Nair, Manoj
; APPLICANT: Chen, Seiyu
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes
; FILE REFERENCE: DEX-0214
; CURRENT APPLICATION NUMBER: US/09/909,567B
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/219,834
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
; LENGTH: 824
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-909-567B-53
```

```

Query Match      8.9%; Score 7; DB 9; Length 824;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 11 AYPVPLR 17
DB 405 AYPVPLR 411
```

```

RESULT 4
US-09-893-519A-36
; Sequence 36, Application US/09893519A
; Publication No. US20030027243A1
; GENERAL INFORMATION:
; APPLICANT: ANADYS PHARMACEUTICALS, INC.
; APPLICANT: THOMPSON, Craig
; APPLICANT: MOORE, Jeffrey
; APPLICANT: BOURMAN, Ed T.
; APPLICANT: BRADLEY, John
; APPLICANT: DESTILVA, Thamara
; APPLICANT: HARRIS, Sandra
; APPLICANT: KOMARNITSKY, Svetlana
; APPLICANT: MENDILLO, Marc
; APPLICANT: MOORE, Daniel
; APPLICANT: MCCOY, Melissa
; APPLICANT: SANDERSON, Karen
; APPLICANT: HAO, Tariq
; APPLICANT: ZHU, Shuhao
; APPLICANT: LONG, Fan
; APPLICANT: DAVIDOV, Eugene
; TITLE OF INVENTION: ANTI-FUNGAL COMPOUNDS AND METHODS OF USE
; FILE REFERENCE: 0342/16548-US2
; CURRENT APPLICATION NUMBER: US/09/893,519A
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,164
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 60/224,457
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 146
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 36
; LENGTH: 880
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Corresponds to SEQ ID NO: 109
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Human Genbank/AAC39727
; DATABASE ENTRY DATE: 1998-05-06
; RELEVANT RESIDUES: (1)..(880)
US-09-893-519A-36
```

```

Query Match      8.9%; Score 7; DB 9; Length 880;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


QY 24 HVLSEFL 30
|111111|
DB 424 HVLSEFL 430

RESULT 5
US-09-746-491-48
; Sequence 48, Application US/09746491
; Patent No. US20020137202A1

GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20020137202A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-621
; CURRENT APPLICATION NUMBER: US/09/746,491
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: USSN 60/171,329
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
; LENGTH: 906
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-746-491-48

Query Match 8.9%; Score 7; DB 10; Length 906;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 VAVNLSR 76
|111111|
DB 117 VAVNLSR 123

RESULT 6
US-09-864-761-45510
; Sequence 45510, Application US/09864761
; Patent No. US20020048763A1

GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45510
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010721.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.82
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.66
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.61
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.63
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.59
; OTHER INFORMATION: EST_HUMAN HIT: BF672312.1, EVALU8 8.00e-04
US-09-864-761-45510

Query Match 7.6%; Score 6; DB 10; Length 30;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 RGPDPD 67
|111111|
DB 12 RGPDPD 17

RESULT 7
US-09-755-109-21
; Sequence 21, Application US/09755109
; Publication No. US20030027981A1

GENERAL INFORMATION:
; APPLICANT: DANO, KEID
; APPLICANT: BLASI, FRANCESCO
; APPLICANT: ROLDAN, ANN LOURING
; APPLICANT: COBELLIS, MARIA VITTORIA
; APPLICANT: MASUCCI, MARIA TERESA
; APPLICANT: APPELLA, ETTORE
; APPLICANT: SCHLEUNING, WOLF-DIETER
; APPLICANT: BEHRENDT, NIELS
; APPLICANT: RONNE, EBBE
; APPLICANT: KRISTENSEN, PETER
; TITLE OF INVENTION: UROKINASE-TYPE PLASMINOGEN ACTIVATOR
; RECEPTOR
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 SEVENTH STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/755,109
; FILING DATE: 08-Jan-2001
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,666

FILING DATE: 22-APR-1996
APPLICATION NUMBER: US 08/319,052
FILING DATE: 06-OCT-1994
APPLICATION NUMBER: US 07/824,189
FILING DATE: 06-DEC-1991
APPLICATION NUMBER: WO PCT/JP90/00090
FILING DATE: 18-OCT-1990
APPLICATION NUMBER: US 07/374,854
FILING DATE: 03-JUL-1989
APPLICATION NUMBER: US 07/334,613
FILING DATE: 07-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, IVER P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: DANO-1P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-755-109-21

Query Match 7.6%; Score 6; DB 9; Length 34;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 OPGPAH 24
|||||
DB 8 OPGPAH 13

RESULT 8
US-09-864-761-41953
Sequence 41953, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO: 41953
LENGTH: 38
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC023344.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.6
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.7
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.9
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.8
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.8
US-09-864-761-41953

Query Match 7.6%; Score 6; DB 10; Length 38;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 47 PLSPTS 52
|||||
DB 24 PLSPTS 29

RESULT 9
US-09-764-887-226
Sequence 226, Application US/09764887
Patent No. US20020042096A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA113
CURRENT APPLICATION NUMBER: US/09/764,887
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 658
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 226
LENGTH: 68
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (6)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-887-226

Query Match 7.6%; Score 6; DB 10; Length 68;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 55 ISPSV 60
|||||
DB 28 ISPSV 33

```

RESULT 10
US-09-864-761-35919
; Sequence 35919, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35919
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC008013.8
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.7
; OTHER INFORMATION: EST_HUMAN HIT: R09295.1, EVALUATE 1.10e+00
US-09-864-761-35919
Query Match 7.6%; Score 6; DB 10; Length 77;
Best Local Similarity 100.0%; Pred. No. 66;

```

```

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 LPLRPA 10
|11111
Db 9 LPLRPA 14

RESULT 11
US-09-764-869-1118
; Sequence 1118, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1118
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (48)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-869-1118
Query Match 7.6%; Score 6; DB 10; Length 86;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 43 LPLAPL 48
|11111
Db 61 LPLAPL 66

RESULT 12
US-09-764-869-798
; Sequence 798, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 798
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (56)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (60)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (79)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (83)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-869-798

```

Query Match 7.6%; Score 6; DB 10; Length 93;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 SVGRGP 64
|||||
DB 37 SVGRGP 42

RESULT 13

US-09-864-761-38224
; Sequence 38224, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38224
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL034384.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.86
; OTHER INFORMATION: SWISSPROT HIT: P08503, EVALU2.70e-01
; OTHER INFORMATION: EST_HUMAN HIT: AV703581.1, EVALU2.10e+00
US-09-864-761-38224

Query Match 7.6%; Score 6; DB 10; Length 114;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 TSVPIS 56
|||||
DB 73 TSVPIS 78

RESULT 14

US-09-771-161A-148
; Sequence 148, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 148
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-771-161A-148

Query Match 7.6%; Score 6; DB 10; Length 131;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 THSLIP 7
|||||
DB 40 THSLIP 45

RESULT 15

US-09-764-847-828
; Sequence 828, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 828
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-847-828

Query Match 7.6%; Score 6; DB 10; Length 147;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LLPRA 10
|||||
DB 108 LLPRA 113

Search completed: March 4, 2003, 12:54:25
Job time: 4.33133 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:36:41 ; Search time 53.8474 Seconds
(without alignments)
1036.856 Million cell updates/sec

Title: US-09-234-208B-2

Perfect score: 419
Sequence: 1 METALCRWGLLALLPPGA.....VGRGPPDAHVAVNLRYEG 419

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size : 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A.Geneseq_101002:*

```
1: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA1980.DAT:*
2: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA1981.DAT:*
3: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA1982.DAT:*
4: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA1983.DAT:*
5: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA1984.DAT:*
6: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA1985.DAT:*
7: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA1986.DAT:*
8: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA1987.DAT:*
9: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA1988.DAT:*
10: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA1989.DAT:*
11: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA1990.DAT:*
12: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA1991.DAT:*
13: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA1992.DAT:*
14: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA1993.DAT:*
15: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA1994.DAT:*
16: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA1995.DAT:*
17: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA1996.DAT:*
18: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA1997.DAT:*
19: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA1998.DAT:*
20: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA1999.DAT:*
21: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA2000.DAT:*
22: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA2001.DAT:*
23: /SIDs2/gcgdata/geneseq/genesqp-emb1/AA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	340	81.1	645	22	Human ErbB2 oncopr
2	340	81.1	645	22	Human ErbB2 extrac
3	340	81.1	653	21	Extracellular HER-
4	340	81.1	653	23	Human Her-2/neu on
5	340	81.1	712	21	Human Her-2/neu fu
6	340	81.1	712	23	Her-2/neu extracel
7	340	81.1	782	18	Her-2-GM-CSF immuno
8	340	81.1	919	21	Human HER-2/neu fu
9	340	81.1	919	23	Her-2/neu extracel
10	340	81.1	1200	21	Human HER-2/neu pr

11	340	81.1	1255	17	AAW01111
12	340	81.1	1255	20	AAW92406
13	340	81.1	1255	21	AAW21198
14	340	81.1	1255	21	AAW84780
15	340	81.1	1255	21	AAW92620
16	340	81.1	1255	22	AAE12130
17	340	81.1	1255	22	AAW85458
18	340	81.1	1255	22	AAW88267
19	340	81.1	1255	22	AAW60167
20	340	81.1	1255	23	AAE24067
21	340	81.1	1255	23	AAE20479
22	340	81.1	1255	23	AAW51143
23	340	81.1	1255	23	AAW77114
24	340	81.1	1255	23	AAW45455
25	340	76.1	951	21	AAW44993
26	307	73.3	1223	23	AAW98923
27	292	69.7	479	22	AAE13112
28	292	69.7	555	22	AAE13108
29	292	69.7	564	22	AAE13110
30	292	69.7	690	22	AAE13109
31	292	69.7	697	22	AAE13111
32	289	69.0	289	22	AAE13120
33	191	45.6	191	19	AAW53354
34	191	45.6	191	20	AAW14572
35	191	45.6	191	22	AAW48767
36	187	44.6	1433	14	AAW39568
37	166	39.6	166	19	AAW59345
38	166	39.6	166	22	AAW62074
39	166	39.6	166	22	AAW48763
40	96	22.9	97	17	AAW00327
41	86	20.5	624	11	AAW08222
42	83	19.8	419	22	AAE09181
43	83	19.8	419	22	AAE09183
44	83	19.8	419	22	AAE09200
45	83	19.8	419	22	AAE09202

ALIGNMENTS

RESULT 1	
AAW60408	standard; Protein; 645 AA.
ID	AAW60408
AC	AAW60408;
DT	24-APR-2001 (first entry)
DE	Human ErbB2 oncoprotein, SEQ ID NO:13.
KW	Anti-ErbB2 monoclonal antibody 2C4; HER2; mouse; murine; humanised; VL;
KW	light chain variable region; cancer; cytostatic; EGFR-expressing cancer;
KW	epidermal growth factor receptor; colon cancer; rectal cancer; tumour;
KW	colorectal cancer; non-small cell lung cancer; metastatic breast cancer;
KW	affinity purification.
OS	Homo sapiens.
PN	WO200100245-A2.
PD	04-JAN-2001.
PF	23-JUN-2000; 2000MO-US17366.
PR	25-JUN-1999; 99US-0141316.
PA	(GETH) GENENTECH INC.
PI	Adams CW, Presta LG, Sliwkowsky M;
DR	WPI; 2001-080862/09.
PT	Treating cancer in a human, where the cancer expresses epidermal growth

PT factor receptor (EGFR), comprises administering an antibody which binds
 PT ErbB2 -
 XX
 PS Example 1; Fig 1A: 89pp; English.

CC The invention relates to a method for treating cancer in a human patient,
 CC wherein the cancer expresses epidermal growth factor receptor (EGFR),
 CC comprising administering an antibody which binds ErbB2 (HER2; AAB60408).
 CC In particular, the anti-ErbB2 antibody is the murine monoclonal antibody
 CC 2C4 (AAB60396, AAB60397) or a humanised version of 2C4 (AAB60398,
 CC AAB60399). The invention also encompasses an isolated nucleic acid
 CC encoding a humanised ErbB2-binding antibody; vectors and host cells
 CC comprising such nucleic acids; the recombinant production of a humanised
 CC ErbB2-binding antibody; and an immunocombinate comprising a humanised
 CC ErbB2-binding antibody and a cytotoxic drug. The ErbB2-binding antibodies
 CC act by antagonising ErbB receptors, and as inhibitors of transforming
 CC growth factor alpha (TGF-alpha)-activated mitogen activated protein
 CC kinase (MAPK). The method of the invention is used for treating cancer,
 CC especially colon cancer, rectal cancer, colorectal cancer, lung cancer
 CC (especially non-small cell lung cancer), or breast cancer (especially
 CC metastatic breast cancer). The antibodies may also have non-therapeutic
 CC uses e.g., as affinity purification agents. Using an antibody which binds
 CC to ErbB2 to treat cancer is preferable to the use of EGFR-targeted
 CC drugs, as EGFR is also highly expressed in other tissues such as the
 CC liver and skin, where the active drug will also bind, with skin toxicity
 CC having been observed for EGFR-targeted drugs. Antibodies which bind
 CC ErbB2 are anticipated to have a better safety profile than such drugs.
 CC The present sequence represents human ErbB2.

XX Sequence 645 AA:

Query Match 81.1%; Score 340; DB 22; Length 645;
 Best Local Similarity 100.0%; Pred. No. 3.9e-302;
 Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 METALCRNGLLIALPPGAASVQVCTGDMKLRPASPEHDMIRHLHYGCGVVGML 60
 DB 1 METALCRNGLLIALPPGAASVQVCTGDMKLRPASPEHDMIRHLHYGCGVVGML 60
 QY 61 ELTYLPTNLSFLDIOEVQGVYLAHNOVQVPLQRIYRGTLQFEDNALAVLDNG 120
 DB 61 ELTYLPTNLSFLDIOEVQGVYLAHNOVQVPLQRIYRGTLQFEDNALAVLDNG 120
 QY 121 DPLNNTPTVGTASPGGLRELQRLSLTEILKGVLIQRPQLCYODTIIMKDIFFHKNNOIA 180
 DB 121 DPLNNTPTVGTASPGGLRELQRLSLTEILKGVLIQRPQLCYODTIIMKDIFFHKNNOIA 180
 QY 181 LTLIDNRSRACHPCSPMKSGRCWGESSESDCSLTRVCAGGCARCKGRLPTDCCHQC 240
 DB 181 LTLIDNRSRACHPCSPMKSGRCWGESSESDCSLTRVCAGGCARCKGRLPTDCCHQC 240
 QY 241 AAGCTPRKSDCLACHFNHSGICELHCPALVYNTDFESMPNPGRTTFGASCTYACP 300
 DB 241 AAGCTPRKSDCLACHFNHSGICELHCPALVYNTDFESMPNPGRTTFGASCTYACP 300
 QY 301 YNLTSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSPCAR 340
 DB 301 YNLTSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSPCAR 340

RESULT 2
 AAB61593
 ID AAB61593 standard; protein; 645 AA.

XX AAB61593;

XX 04-APR-2001 (first entry)

XX Human ErbB2 extracellular domain.

XX Human; ErbB2; cytosolic; prostate cancer; receptor tyrosine kinase;
 KW antibody; ErbB receptor; monoclonal antibody 2C4; variable light chain.
 XX

OS Homo sapiens.

XX MO200100238-A1.

XX 04-JAN-2001.

XX 23-JUN-2000; 2000WO-US17423.

XX 25-JUN-1999; 99US-0141315.

XX (GENTH) GENENTECH INC.

XX (SLOK) SLOAN KETTERING INST CANCER RES.

XX Agus DB, Scher HI, Sliwkowski MX;

XX WPI: 2001-159131/16.

XX Treating prostate cancer in a human comprises administering an antibody

XX which binds ErbB2 and blocks ligand activation of an ErbB receptor -

XX Disclosure: Fig 1: 93pp; English.

CC The ErbB family of receptor tyrosine kinases are important mediators of
 CC cell growth, differentiation and survival. The receptor family includes
 CC four distinct members including Epidermal Growth Factor Receptor (EGFR or
 CC ErbB1), HER2 (ErbB2 or p185^{neu}), HER3 (ErbB3) and Her4 (ErbB4 or tyro2).
 CC The present invention relates to a method for treating prostate cancer.
 CC The method comprises administering an antibody which binds ErbB2 and
 CC blocks ligand activation of an ErbB receptor. Preferably, the antibody
 CC blocks binding of monoclonal antibody 2C4 to ErbB2 and/or blocks
 CC TGF-alpha activation of mitogen-activated protein kinase (MAPK). The
 CC present sequence is the extracellular domain of human ErbB2.

XX Sequence 645 AA:

Query Match 81.1%; Score 340; DB 22; Length 645;
 Best Local Similarity 100.0%; Pred. No. 3.9e-302;
 Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 METALCRNGLLIALPPGAASVQVCTGDMKLRPASPEHDMIRHLHYGCGVVGML 60
 DB 1 METALCRNGLLIALPPGAASVQVCTGDMKLRPASPEHDMIRHLHYGCGVVGML 60
 QY 61 ELTYLPTNLSFLDIOEVQGVYLAHNOVQVPLQRIYRGTLQFEDNALAVLDNG 120
 DB 61 ELTYLPTNLSFLDIOEVQGVYLAHNOVQVPLQRIYRGTLQFEDNALAVLDNG 120
 QY 121 DPLNNTPTVGTASPGGLRELQRLSLTEILKGVLIQRPQLCYODTIIMKDIFFHKNNOIA 180
 DB 121 DPLNNTPTVGTASPGGLRELQRLSLTEILKGVLIQRPQLCYODTIIMKDIFFHKNNOIA 180
 QY 181 LTLIDNRSRACHPCSPMKSGRCWGESSESDCSLTRVCAGGCARCKGRLPTDCCHQC 240
 DB 181 LTLIDNRSRACHPCSPMKSGRCWGESSESDCSLTRVCAGGCARCKGRLPTDCCHQC 240
 QY 241 AAGCTPRKSDCLACHFNHSGICELHCPALVYNTDFESMPNPGRTTFGASCTYACP 300
 DB 241 AAGCTPRKSDCLACHFNHSGICELHCPALVYNTDFESMPNPGRTTFGASCTYACP 300
 QY 301 YNLTSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSPCAR 340
 DB 301 YNLTSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSPCAR 340

RESULT 3
 AAB21200
 ID AAB21200 standard; protein; 653 AA.

XX AAB21200;

XX 12-JAN-2001 (first entry)

XX Extracellular HER-2/neu protein.


```
XX HER-2/neu; oncogene; tyrosine kinase; cytosolic; vaccine;  
KW breast cancer; prostate cancer; ovarian cancer; lung cancer;  
KM colon cancer.  
XX  
XX  
OS Unidentified.  
PN WO200044899-A1.  
PD 03-AUG-2000.  
XX  
XX 28-JAN-2000; 2000WO-US02164.  
PF 29-JAN-1999; 99GS-0117976.  
XX  
XX (CORI-) CORIXA CORP.  
PA (SMK) SMITHKLINE BEECHAM.  
XX  
XX Cheever MA, Gheysen D;  
PI  
XX WPI: 2000-505976/45.  
DR  
XX  
XX HER-2/neu extracellular domain/phosphorylation domain fusion proteins  
PT useful for vaccinating against breast, ovarian, colon, lung and  
PT prostate cancers -  
XX  
XX  
XX Claim 2; Fig 9; 128pp; English.  
XX  
XX The present sequence is the extracellular HER-2/neu protein. HER-2/neu is  
CC a member of the tyrosine kinase family of receptor-like glycoproteins and  
CC shows homology to the epidermal growth factor receptor (EGFR). It  
CC probably plays a part in cell growth and/or differentiation. The  
CC HER-2/neu gene is an oncogene. An HER-2/neu fusion protein comprising a  
CC HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation  
CC domain may be used to treat or prevent cancer by eliciting or  
CC enhancing an immune response to the HER-2/neu protein. It may be used  
CC to treat malignancies such as breast, ovarian, colon, lung and  
CC prostate cancers, and may be used as an antigen to vaccinate against  
CC these neoplasias.  
CC  
CC  
SQ Sequence 653 AA:  
  
Query Match 81.1%; Score 340; DB 21; Length 653;  
Best Local Similarity 100.0%; Pred. No. 4e-302;  
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MELAALCRWGLLALLPFGAASVCTGDMKRLPASPETHLMDRLHYOGCQVVOGNTL 60  
DB 1 MELAALCRWGLLALLPFGAASVCTGDMKRLPASPETHLMDRLHYOGCQVVOGNTL 60  
QY 61 ELTYLPTNASTSLFLADIDQEVQGVYLAHNVQVPLQRLRIYRGTLDFEDNTALAVLDNG 120  
DB 61 ELTYLPTNASTSLFLADIDQEVQGVYLAHNVQVPLQRLRIYRGTLDFEDNTALAVLDNG 120  
QY 121 DELNNTPTVGTASPGSLRELOLRSLTEILKGVLTDRNQLCYQDTIILKKDIFHKNNQIA 180  
DB 121 DELNNTPTVGTASPGSLRELOLRSLTEILKGVLTDRNQLCYQDTIILKKDIFHKNNQIA 180  
QY 181 LTLIDTNRSRACHPCSPMKSGSRCSWESSESDCSILRTYCAGACARCKPLPTDCCHQC 240  
DB 181 LTLIDTNRSRACHPCSPMKSGSRCSWESSESDCSILRTYCAGACARCKPLPTDCCHQC 240  
QY 241 AAGCTGPKHSDDLACHLHFHNSGICELHCPALVYNTDFESMPNPEGRYTFGASCVTACP 300  
DB 241 AAGCTGPKHSDDLACHLHFHNSGICELHCPALVYNTDFESMPNPEGRYTFGASCVTACP 300  
QY 301 YNVLSTDVGSCTLVCPHLHNOEYTAEDGTORCEKCSKPCAR 340  
DB 301 YNVLSTDVGSCTLVCPHLHNOEYTAEDGTORCEKCSKPCAR 340
```

```
ID AAM51145 standard; Protein; 653 AA.  
XX  
XX AAM51145;  
AC  
XX  
DT 17-JUN-2002 (first entry)  
XX  
XX Human Her-2/neu oncoprotein extracellular domain.  
XX  
XX Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;  
KM tyrosine kinase; receptor; c-erbB2; gene therapy.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200212341-A2.  
PN  
XX  
XX 14-FEB-2002.  
PD  
XX  
XX 03-AUG-2001; 2001WO-US24283.  
PF  
XX  
XX 03-AUG-2000; 2000US-0632507.  
PR  
XX  
XX (CORI-) CORIXA CORP.  
PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.  
XX  
XX Cheever MA, Gheysen D;  
PI  
XX  
XX WPI: 2002-241743/29.  
DR  
XX  
XX Her-2/neu fusion protein for treating or preventing cancer by eliciting  
PT or enhancing an immune response to the protein, has Her-2/neu  
PT extracellular domain fused to Her-2/neu intracellular or  
PT phosphorylation domain -  
XX  
XX  
XX Claim 2; Fig 9; 141pp; English.  
XX  
XX The present sequence is that of the extracellular domain of  
CC human Her-2/neu (p185 glycoprotein or c-erbB2), an oncogenic  
CC self-protein and target for anti-cancer vaccines. The Her-2/neu  
CC gene is amplified and p185 is overexpressed in a variety of cancers,  
CC including breast, ovarian, colon, lung and prostate cancer.  
CC Her-2/neu (see AAM51143) is a member of the tyrosine kinase family  
CC of receptor-like glycoproteins. It comprises an extracellular  
CC domain with homology to the epidermal growth factor receptor  
CC (EGFR), a highly hydrophobic transmembrane domain and a C-terminal  
CC intracellular domain that also shows homology to EGFR. Its  
CC overexpression correlates with a poor prognosis in breast and  
CC ovarian cancers. The invention provides Her-2/neu fusion  
CC proteins, nucleic acids encoding them, viral vectors, and vaccines  
CC comprising the fusion proteins or nucleic acid molecules. In  
CC preferred fusion proteins, the extracellular domain of a Her-2/neu  
CC protein is fused to a Her-2/neu intracellular domain or  
CC phosphorylation domain (or its deltaCD fragment). An immune  
CC response to Her-2/neu protein is elicited or enhanced by  
CC administering the fusion protein in the form of a vaccine, or by  
CC transfecting cells of an animal ex vivo with a nucleic acid  
CC encoding the fusion protein, and delivering the transfected cells  
CC to the animal. The fusion proteins, nucleic acids, and isolated  
CC specific T-cells are useful for inhibiting the development of a  
CC cancer, especially breast, ovarian, colon, lung or prostate cancer  
CC in a patient. T cells that specifically react with a Her-2/neu  
CC fusion protein can be used to remove tumour cells from a sample in  
CC order to inhibit the development of cancer in a patient.  
CC  
CC  
SQ Sequence 653 AA:  
  
Query Match 81.1%; Score 340; DB 23; Length 653;  
Best Local Similarity 100.0%; Pred. No. 4e-302;  
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MELAALCRWGLLALLPFGAASVCTGDMKRLPASPETHLMDRLHYOGCQVVOGNTL 60  
DB 1 MELAALCRWGLLALLPFGAASVCTGDMKRLPASPETHLMDRLHYOGCQVVOGNTL 60
```

QY 61 ELTYLPTNALSFLQDIQEVGVYLIANQVROPVLRIRIVRGTLQFEDNYALAVLNG 120
 |||||
 Db 61 ELTYLPTNALSFLQDIQEVGVYLIANQVROPVLRIRIVRGTLQFEDNYALAVLNG 120
 QY 121 DPLNNTPTVGTASPGGLRELQLSLTELKGVLIORNPOLCYODTILMKDIFHKNNQLA 180
 |||||
 Db 121 DPLNNTPTVGTASPGGLRELQLSLTELKGVLIORNPOLCYODTILMKDIFHKNNQLA 180
 QY 181 LTLIDTNRSRACHPCSPMKGSKRCWGESSEDDOSLTRVCAGGCARCKGPLPTDCCHEOC 240
 |||||
 Db 181 LTLIDTNRSRACHPCSPMKGSKRCWGESSEDDOSLTRVCAGGCARCKGPLPTDCCHEOC 240
 QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFESMPNPEGRTYFGASCVTACP 300
 |||||
 Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFESMPNPEGRTYFGASCVTACP 300
 QY 301 YNLTSTDVGSCTLVCPHNOEVTAEEDGTORCEKSKPCAR 340
 |||||
 Db 301 YNLTSTDVGSCTLVCPHNOEVTAEEDGTORCEKSKPCAR 340

RESULT 5
 AAB21204
 ID AAB21204 standard; protein; 712 AA.
 AC AAB21204;
 XX
 XX
 DT 12-JAN-2001 (first entry)
 DE Human HER-2/neu fusion protein.
 XX
 XX Human; HER-2/neu; oncogene; tyrosine kinase; cytosolic; vaccine;
 KW breast cancer; prostate cancer; ovarian cancer; lung cancer;
 KW colon cancer; fusion protein.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200044899-A1.
 XX
 PD 03-AUG-2000.
 XX
 PF 28-JAN-2000; 2000WO-US02164.
 XX
 PR 29-JAN-1999; 99US-0117976.
 XX
 PA (CORI-) CORIXA CORP.
 PA (SMK) SMITHKLINE BEECHAM.
 XX
 PI Cheever MA, Gheysen D;
 XX
 DR WPI: 2000-505976/45.
 XX
 PT HER-2/neu extracellular domain/phosphorylation domain fusion proteins
 PT useful for vaccinating against breast, ovarian, colon, lung and
 PT prostate cancers -
 XX
 PS Claim 27; Fig 13; 128pp; English.
 XX
 CC The present sequence is a fusion protein comprising the extracellular
 CC domain and a preferred portion of the phosphorylation domain of the human
 CC HER-2/neu protein. HER-2/neu is a member of the tyrosine kinase family of
 CC receptor-like glycoproteins and shows homology to the epidermal growth
 CC factor receptor (EGFR). It probably plays a part in cell growth and/or
 CC differentiation. The HER-2/neu gene is an oncogene. HER-2/neu fusion
 CC proteins may be used to treat or prevent cancer by eliciting or enhancing
 CC an immune response to the HER-2/neu protein. They may be used to treat
 CC malignancies such as breast, ovarian, colon, lung and prostate cancers,
 CC and may be used as an antigen to vaccinate against these neoplasias.
 XX
 PS Sequence 712 AA:
 QY Query Match 81.1%; Score 340; DB 21; Length 712;

Best Local Similarity 100.0%; Pred. No. 4,3e-302;
 Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MELAALCRWGLLALIPPAASTOVCTGDMKRLPASPETHLDMIRHLXGCCOVVQNL 60
 |||||
 Db 1 MELAALCRWGLLALIPPAASTOVCTGDMKRLPASPETHLDMIRHLXGCCOVVQNL 60
 QY 61 ELTYLPTNALSFLQDIQEVGVYLIANQVROPVLRIRIVRGTLQFEDNYALAVLNG 120
 |||||
 Db 61 ELTYLPTNALSFLQDIQEVGVYLIANQVROPVLRIRIVRGTLQFEDNYALAVLNG 120
 QY 121 DPLNNTPTVGTASPGGLRELQLSLTELKGVLIORNPOLCYODTILMKDIFHKNNQLA 180
 |||||
 Db 121 DPLNNTPTVGTASPGGLRELQLSLTELKGVLIORNPOLCYODTILMKDIFHKNNQLA 180
 QY 181 LTLIDTNRSRACHPCSPMKGSKRCWGESSEDDOSLTRVCAGGCARCKGPLPTDCCHEOC 240
 |||||
 Db 181 LTLIDTNRSRACHPCSPMKGSKRCWGESSEDDOSLTRVCAGGCARCKGPLPTDCCHEOC 240
 QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFESMPNPEGRTYFGASCVTACP 300
 |||||
 Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFESMPNPEGRTYFGASCVTACP 300
 QY 301 YNLTSTDVGSCTLVCPHNOEVTAEEDGTORCEKSKPCAR 340
 |||||
 Db 301 YNLTSTDVGSCTLVCPHNOEVTAEEDGTORCEKSKPCAR 340

RESULT 6
 AAM51149
 ID AAM51149 standard; Protein; 712 AA.
 AC AAM51149;
 XX
 XX
 DT 17-JUN-2002 (first entry)
 DE Her-2/neu extracellular domain-delta-phosphorylation domain fusion.
 XX
 XX Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;
 KW tyrosine kinase; receptor; c-erbB2; gene therapy.
 XX
 OS Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT Domain 1..653
 FT /note= "extracellular domain"
 FT Domain 654..712
 FT /note= "phosphorylation domain fragment"
 XX
 PN WO200212341-A2.
 XX
 PD 14-FEB-2002.
 XX
 PF 03-AUG-2001; 2001WO-US24283.
 XX
 PR 03-AUG-2000; 2000US-0632507.
 XX
 PA (CORI-) CORIXA CORP.
 PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 PI Cheever MA, Gheysen D;
 XX
 DR WPI: 2002-241743/29.
 XX
 PT Her-2/neu fusion protein for treating or preventing cancer by eliciting
 PT or enhancing an immune response to the protein, has Her-2/neu
 PT extracellular domain fused to Her-2/neu intracellular or
 PT phosphorylation domain -
 XX
 PS Claim 37; Fig 13; 141pp; English.
 XX
 CC The present sequence is that of a fusion protein between the
 CC extracellular domain and a fragment (Delta) of the phosphorylation

CC domain of human Her-2/neu (see AAM5143), an oncogenic self-protein
CC and target for anti-cancer vaccines. The fusion protein can be
CC obtained by recombinant DNA methods. Her-2/neu overexpression
CC correlates with a poor prognosis in breast and ovarian cancers.
CC The invention provides Her-2/neu fusion proteins, nucleic acids
CC encoding them, viral vectors, and vaccines comprising the fusion
CC proteins or nucleic acid molecules. In preferred fusion proteins,
CC the extracellular domain of Her-2/neu is fused to a Her-2/neu
CC intracellular domain or phosphorylation domain (or its Deltap
CC fragment). An immune response to Her-2/neu protein is elicited or
CC enhanced by administering the fusion protein in the form of a vaccine,
CC or by transfecting cells of an animal ex vivo with a nucleic acid
CC encoding the fusion protein, and delivering the transfected cells
CC to the animal. The fusion proteins, nucleic acids, and isolated
CC specific T-cells are useful for inhibiting the development of a
CC cancer, especially breast, ovarian, colon, lung or prostate cancer
CC in a patient. T cells that specifically react with a Her-2/neu
CC fusion protein can be used to remove tumour cells from a sample in
CC order to inhibit the development of cancer in a patient.

SO Sequence 712 AA:

Query Match 81.1%; Score 340; DB 23; Length 712;
Best Local Similarity 100.0%; Pred. No. 4.3e-302;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 METALACRMGLLLALLPRGAASQVCTGTDMLRLPASPEITHDMLRHLYQGCGVVGNL 60
DB 1 METALACRMGLLLALLPRGAASQVCTGTDMLRLPASPEITHDMLRHLYQGCGVVGNL 60
OY 61 ELTYLPTNASLSFLQDIOEVQGYVLLAHNOVROVPLQRLRIVRGTOLEFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIOEVQGYVLLAHNOVROVPLQRLRIVRGTOLEFEDNYALAVLDNG 120
OY 121 DPLNNTPTVTGASPGGRLREQLRSLEIFELKGVLIQRNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVTGASPGGRLREQLRSLEIFELKGVLIQRNPOLCYODTILMKDIFHKNNOLA 180
OY 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDCOSLRTVAGGCARCKGPLPDDCHEQC 240
DB 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDCOSLRTVAGGCARCKGPLPDDCHEQC 240
OY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYTACP 300
OY 301 YNYLSTDVGSCTLVCPPLHNOEVTAEADGTORCEKCSKPCAR 340
DB 301 YNYLSTDVGSCTLVCPPLHNOEVTAEADGTORCEKCSKPCAR 340

RESULT 7

ID AAM19764
AAW19764 standard; Protein; 782 AA.

XX AAW19764;

DT 17-SEP-1997 (first entry)

DE Her2-GM-CSF immunostimulant fusion protein.

KM Her2-GM-CSF; granulocyte macrophage colony stimulating factor;

KW growth factor receptor; oncogene; immunostimulant; cancer;

XX Homo sapiens.

OS Homo sapiens.

FH Key

FT Protein

FT Peptide

FT Protein

DT Protein

/label= GM-CSF

W09724438-A1.

10-JUL-1997.

23-DEC-1996; 96WO-US20241.

28-DEC-1995; 95US-0579823.

(ACTI-) ACTIVATED CELL THERAPY INC.

Laus R, Ruegg CL, Wu H;

WPI: 1997-363674/33.

N-PSDB; AAT72725.

Potent APC that activates T-cells to give multivalent cellular

immune response - can also induce a cytotoxic T-cell response in a

vertebrate subject

Disclosure: Fig 8; 45pp; English.

A fusion protein (AAW19764) comprises Her2 (a growth factor receptor

that is overexpressed in breast, ovarian can other cancer cells)

and granulocyte-macrophage colony stimulating factor (GM-CSF). It

is the expression product of a nucleic acid molecule (AAT72725)

prep. by PCR amplification of Her2 cDNA from a breast cancer cell

line and fusion to GM-CSF cDNA. Fusion expression vectors can be

used to transfect mammalian and insect cells. The Her2-GM-CSF

fusion protein is used to generate anti-Her2 immunity. Tumour

cells are eliminated by cytotoxic T lymphocytes activated in vivo

or in vitro by exposure to antigen-presenting cells exposed to the

fusion protein.

SQ Sequence 782 AA:

Query Match 81.1%; Score 340; DB 18; Length 782;
Best Local Similarity 100.0%; Pred. No. 4.6e-302;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 METALACRMGLLLALLPRGAASQVCTGTDMLRLPASPEITHDMLRHLYQGCGVVGNL 60
DB 1 METALACRMGLLLALLPRGAASQVCTGTDMLRLPASPEITHDMLRHLYQGCGVVGNL 60
OY 61 ELTYLPTNASLSFLQDIOEVQGYVLLAHNOVROVPLQRLRIVRGTOLEFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIOEVQGYVLLAHNOVROVPLQRLRIVRGTOLEFEDNYALAVLDNG 120
OY 121 DPLNNTPTVTGASPGGRLREQLRSLEIFELKGVLIQRNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVTGASPGGRLREQLRSLEIFELKGVLIQRNPOLCYODTILMKDIFHKNNOLA 180
OY 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDCOSLRTVAGGCARCKGPLPDDCHEQC 240
DB 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDCOSLRTVAGGCARCKGPLPDDCHEQC 240
OY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYTACP 300
OY 301 YNYLSTDVGSCTLVCPPLHNOEVTAEADGTORCEKCSKPCAR 340
DB 301 YNYLSTDVGSCTLVCPPLHNOEVTAEADGTORCEKCSKPCAR 340

RESULT 8

ID AAB21203
AAB21203 standard; protein; 919 AA.

XX AAB21203;

DT 12-JAN-2001 (first entry)

```

XX  Human HER-2/neu fusion protein.
DE
XX
XX  Human: HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;
KM breast cancer; prostate cancer; ovarian cancer; lung cancer;
KW colon cancer; fusion protein.
XX
XX  Homo sapiens.
OS Synthetic.
XX
XX  WO200044899-A1.
PN
XX
XX  03-AUG-2000.
PD
XX
XX  28-JAN-2000; 2000WO-00502164.
PF
XX
XX  29-JAN-1999; 99US-0117976.
PR
XX
XX  (CORI-) CORIXA CORP.
PA (SMK ) SMITHKLINE BEECHAM.
XX
XX  Cheever MA, Gheysen D;
PI
XX
XX  WPI; 2000-505976/45.
DR
XX
XX  HER-2/neu extracellular domain/phosphorylation domain fusion proteins
PT useful for vaccinating against breast, ovarian, colon, lung and
PT prostate cancers -
XX
XX  Claim 2: Fig 12: 128pp; English.
PS
XX
XX  The present sequence is a fusion protein comprising the extracellular
CC domain and the phosphorylation domain of the human HER-2/neu protein.
CC HER-2/neu is a member of the tyrosine kinase family of receptor-like
CC glycoproteins and shows homology to the epidermal growth factor receptor
CC (EGFR). It probably plays a part in cell growth and/or differentiation.
CC The HER-2/neu gene is an oncogene. HER-2/neu fusion proteins may be used
CC to treat or prevent cancer by eliciting or enhancing an immune response
CC to the HER-2/neu protein. They may be used to treat malignancies such as
CC breast, ovarian, colon, lung and prostate cancers, and may be used as an
CC antigen to vaccinate against these neoplasias.
XX
XX  Sequence 919 AA:
SQ
Query Match 81.1%; Score 340; DB 21; Length 919;
Best Local Similarity 100.0%; Pred. No. 5.3e-302;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
0;
OY 1 METALCRMGLLLALLPPGAASVQVCTGTDMLRLPASPTHLDMLRHLYOGCGVQGNL 60
DB 1 METALCRMGLLLALLPPGAASVQVCTGTDMLRLPASPTHLDMLRHLYOGCGVQGNL 60
OY 61 ELTYLPTNALSFLQDQEVGYVLLAHNOVROPRLRLRIVRTQLFEENYALAVLDNG 120
DB 61 ELTYLPTNALSFLQDQEVGYVLLAHNOVROPRLRLRIVRTQLFEENYALAVLDNG 120
OY 121 DPLNNTTPVAGSPGGLREVLQSLRSLTEILKGVLIQRNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTTPVAGSPGGLREVLQSLRSLTEILKGVLIQRNPOLCYODTILMKDIFHKNNOLA 180
OY 121 DPLNNTTPVAGSPGGLREVLQSLRSLTEILKGVLIQRNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTTPVAGSPGGLREVLQSLRSLTEILKGVLIQRNPOLCYODTILMKDIFHKNNOLA 180
OY 181 LTLIDTNRSPACHPCSPMCKSGKRWGESSSDCSLRTVACGAGCARCKGPLPTDCHEQC 240
DB 181 LTLIDTNRSPACHPCSPMCKSGKRWGESSSDCSLRTVACGAGCARCKGPLPTDCHEQC 240
OY 241 AAGCTGKHSDDCLACHFNISGICELHCPALVYNTDTPFSMPERRYFGASCVTACP 300
DB 241 AAGCTGKHSDDCLACHFNISGICELHCPALVYNTDTPFSMPERRYFGASCVTACP 300
OY 301 YNYLSTDVSGCTLVCPPLHNOEVTAAEDGTORCEKSKPCAR 340
DB 301 YNYLSTDVSGCTLVCPPLHNOEVTAAEDGTORCEKSKPCAR 340

```

```

RESULT 9
ID AAM51148
XX AAM51148 standard; Protein; 919 AA.
XX
XX AAM51148;
AC
XX
XX 17-JUN-2002 (first entry)
DT
XX
XX Her-2/neu extracellular domain-phosphorylation domain fusion.
DE
XX
XX Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;
KM tyrosine kinase; receptor; c-erbB2; gene therapy.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Domain 1..653
FT Domain /note="extracellular domain"
FT Domain 654..919
FT Domain /note="phosphorylation domain"
XX
XX WO200212341-A2.
PN
XX
XX 14-FEB-2002.
PD
XX
XX 03-AUG-2001; 2001WO-US24283.
PF
XX
XX 03-AUG-2000; 2000US-0632507.
PR
XX
XX (CORI-) CORIXA CORP.
PA (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Cheever MA, Gheysen D;
PI
XX
XX WPI; 2002-241743/29.
DR
XX
XX Her-2/neu fusion protein for treating or preventing cancer by eliciting
PT or enhancing an immune response to the protein, has Her-2/neu
PT extracellular domain fused to Her-2/neu intracellular or
PT phosphorylation domain -
XX
XX
XX Claim 2: Fig 12: 141pp; English.
PS
XX
XX The present sequence is that of a fusion protein between the
CC extracellular domain and phosphorylation domain of human Her-2/neu
CC (see AAM51143), an oncogenic self-protein and target for anti-cancer
CC vaccines. The fusion protein can be obtained by recombinant DNA
CC methods. Her-2/neu overexpression correlates with a poor prognosis
CC in breast and ovarian cancers. The invention provides Her-2/neu
CC fusion proteins, nucleic acids encoding them, viral vectors, and
CC vaccines comprising the fusion proteins or nucleic acid molecules.
CC In preferred fusion proteins, the extracellular domain of a
CC Her-2/neu protein is fused to a Her-2/neu intracellular domain or
CC phosphorylation domain (or its deltaap fragment). An immune
CC response to Her-2/neu protein is elicited or enhanced by
CC administering the fusion protein in the form of a vaccine, or by
CC transfecting cells of an animal ex vivo with a nucleic acid
CC encoding the fusion protein, and delivering the transfected cells
CC to the animal. The fusion proteins, nucleic acids, and isolated
CC specific T-cells are useful for inhibiting the development of a
CC cancer, especially breast, ovarian, colon, lung or prostate cancer
CC in a patient. T cells that specifically react with a Her-2/neu
CC fusion protein can be used to remove tumour cells from a sample in
CC order to inhibit the development of cancer in a patient.
XX
XX Sequence 919 AA:
SQ
Query Match 81.1%; Score 340; DB 23; Length 919;
Best Local Similarity 100.0%; Pred. No. 5.3e-302;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
0;
OY 1 METALCRMGLLLALLPPGAASVQVCTGTDMLRLPASPTHLDMLRHLYOGCGVQGNL 60
DB 1 METALCRMGLLLALLPPGAASVQVCTGTDMLRLPASPTHLDMLRHLYOGCGVQGNL 60

```


CC Human HER-2/neu protein (AAW0111), also called p185 or c-erbB2, is
 CC the product of the HER-2/neu oncogene (see also AA040739). The
 CC protein is over-expressed in various cancers, including breast,
 CC ovarian, colon, lung and prostate. The intracellular domain of the
 CC protein can be used to immunise an animal against a malignancy with
 CC which the oncogene is associated. The polypeptide can be produced
 CC in transformed host cells for use in immunisation. Alternatively,
 CC animal cells are transfected in vivo or ex vivo with a viral vector
 CC that directs expression of the polypeptide.

XX Sequence 1255 AA:

Query Match 81.1%; Score 340; DB 17; Length 1255;

Best Local Similarity 100.0%; Pred. No. 6.9e-302; Mismatches 0; Indels 0; Gaps 0;

Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELAALCRMGILLALLPPGAASVQVCTGTDMLRLPASPEHIDMLRHLYQGQVVGNL 60

DB 1 MELAALCRMGILLALLPPGAASVQVCTGTDMLRLPASPEHIDMLRHLYQGQVVGNL 60

QY 61 ELTYLPTNASLSFLQDIOEVQGYVLIANNOVROPPLQRLRIVGTQLFEDNVALAVLDNG 120

DB 61 ELTYLPTNASLSFLQDIOEVQGYVLIANNOVROPPLQRLRIVGTQLFEDNVALAVLDNG 120

QY 121 DPLNNTTPVTVGASPGGLRELQLSLFEILKGVLIQRNPOLCYQDTILMKDIFHKNNOLA 180

DB 121 DPLNNTTPVTVGASPGGLRELQLSLFEILKGVLIQRNPOLCYQDTILMKDIFHKNNOLA 180

QY 181 LTLIDNRSRACHPCSPMCKSGSRGWESSSDCSLFTVACAGCARGKGLPTDCHEQC 240

DB 181 LTLIDNRSRACHPCSPMCKSGSRGWESSSDCSLFTVACAGCARGKGLPTDCHEQC 240

QY 241 AACGTGPKHSDCLACHFNHSGICELHCPALVYNTDFFESMPNPEGRYTFGASCVTACP 300

DB 241 AACGTGPKHSDCLACHFNHSGICELHCPALVYNTDFFESMPNPEGRYTFGASCVTACP 300

QY 301 YNYLSTDVGSCTLVCPPLHNOEYTAEDGTORCEKSKPCAR 340

DB 301 YNYLSTDVGSCTLVCPPLHNOEYTAEDGTORCEKSKPCAR 340

RESULT 12

AAW92406 ID AAW92406 standard; Protein; 1255 AA.

XX AAW92406;

XX 21-APR-1999 (first entry)

XX Human HER-2/neu oncogene protein.

XX HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;

XX malignancy; treatment; tumour.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Region 676..1255

XX /note="region which elicits immune response"

XX US5869445-A.

XX 09-FEB-1999.

XX 01-APR-1996; 9605-0625101.

XX 01-APR-1996; 9605-0625101.

XX 17-MAR-1993; 9305-0033644.

XX 12-AUG-1993; 9305-0106112.

XX 31-MAR-1995; 9505-0414417.

XX (UNIW) UNIV WASHINGTON.

PI Cheever MA, Disis ML;

XX WPI; 1999-152835/13.

XX N-PSDB; AAX01912.

XX Use of HER-2/neu polypeptides - for eliciting an immune response to

XX an HER-2/neu associated malignancy, particularly for treating or

XX preventing tumours

XX Claim 3; Column 31-38; 26pp; English.

PS This sequence represents the human HER-2/neu oncogene protein. A fragment
 CC of this protein is used in a method for eliciting or enhancing an immune
 CC response to HER-2/neu protein. The polypeptide can stimulate T cells and
 CC B cells to produce an immune response to the HER-2/neu protein. The
 CC method can be used for immunisation against a malignancy in which the
 CC HER-2/neu oncogene is associated and in the treatment of an existing
 CC tumour, or to prevent tumour occurrence or recurrence.

XX Sequence 1255 AA:

Query Match 81.1%; Score 340; DB 20; Length 1255;

Best Local Similarity 100.0%; Pred. No. 6.9e-302; Mismatches 0; Indels 0; Gaps 0;

Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELAALCRMGILLALLPPGAASVQVCTGTDMLRLPASPEHIDMLRHLYQGQVVGNL 60

DB 1 MELAALCRMGILLALLPPGAASVQVCTGTDMLRLPASPEHIDMLRHLYQGQVVGNL 60

QY 61 ELTYLPTNASLSFLQDIOEVQGYVLIANNOVROPPLQRLRIVGTQLFEDNVALAVLDNG 120

DB 61 ELTYLPTNASLSFLQDIOEVQGYVLIANNOVROPPLQRLRIVGTQLFEDNVALAVLDNG 120

QY 121 DPLNNTTPVTVGASPGGLRELQLSLFEILKGVLIQRNPOLCYQDTILMKDIFHKNNOLA 180

DB 121 DPLNNTTPVTVGASPGGLRELQLSLFEILKGVLIQRNPOLCYQDTILMKDIFHKNNOLA 180

QY 181 LTLIDNRSRACHPCSPMCKSGSRGWESSSDCSLFTVACAGCARGKGLPTDCHEQC 240

DB 181 LTLIDNRSRACHPCSPMCKSGSRGWESSSDCSLFTVACAGCARGKGLPTDCHEQC 240

QY 241 AACGTGPKHSDCLACHFNHSGICELHCPALVYNTDFFESMPNPEGRYTFGASCVTACP 300

DB 241 AACGTGPKHSDCLACHFNHSGICELHCPALVYNTDFFESMPNPEGRYTFGASCVTACP 300

QY 301 YNYLSTDVGSCTLVCPPLHNOEYTAEDGTORCEKSKPCAR 340

DB 301 YNYLSTDVGSCTLVCPPLHNOEYTAEDGTORCEKSKPCAR 340

RESULT 13

AAB21198 ID AAB21198 standard; protein; 1255 AA.

XX AAB21198;

XX 12-JAN-2001 (first entry)

XX Human HER-2/neu protein..

XX Human; HER-2/neu; oncogene; tyrosine kinase; cytosolic; vaccine;

XX breast cancer; prostate cancer; ovarian cancer; lung cancer;

XX colon cancer.

XX Homo sapiens.

XX WO200044899-A1.

XX 03-AUG-2000.

XX 28-JAN-2000; 2000WO-US02164.

XX 29-JAN-1999; 9905-0117976.

```

XX (CORI-) CORIXA CORP.
PA (SMIX ) SMITHLINE BEECHAM.
XX
XX Cheever MA, Gheysen D;
PI
XX WPI: 2000-505976/45.
DR N-PSDB: AAA89736.
XX
PT HER-2/neu extracellular domain/phosphorylation domain fusion proteins
PT useful for vaccinating against breast, ovarian, colon, lung and
PT prostate cancers -
XX
PS Claim 52; Fig 7; 128pp; English.
XX
CC The present sequence is the human HER-2/neu protein. It is a member of
CC the tyrosine kinase family of receptor-like glycoproteins and shows
CC homology to the epidermal growth factor receptor (EGFR). It probably
CC plays a part in cell growth and/or differentiation. The HER-2/neu
CC gene is an oncogene. An HER-2/neu fusion protein comprising a
CC HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation
CC domain may be used to treat or prevent cancer by eliciting or
CC enhancing an immune response to the HER-2/neu protein. It may be used
CC to treat malignancies such as breast, ovarian, colon, lung and
CC prostate cancers, and may be used as an antigen to vaccinate against
CC these neoplasias.
XX
SQ Sequence 1255 AA;
XX
XX Query Match 81.1%; Score 340; DB 21; Length 1255;
XX Best Local Similarity 100.0%; Pred. No. 6.9e-302;
XX Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MELAAACRNGLLALLPPAASSTQVCTGDMKRLRPASETHLDMRLHXGCGVQGNL 60
XX |
XX 1 MELAAACRNGLLALLPPAASSTQVCTGDMKRLRPASETHLDMRLHXGCGVQGNL 60
DB
QY 61 ELTYLPTNASLFLDIQIEVQGYVLLAHNOVROVPLQRLRIYVGTQLFEDNVALAVLDNG 120
XX |
XX 61 ELTYLPTNASLFLDIQIEVQGYVLLAHNOVROVPLQRLRIYVGTQLFEDNVALAVLDNG 120
DB
QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQRPOLCYQDTILMKDIFHKNNOLA 180
XX |
XX 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQRPOLCYQDTILMKDIFHKNNOLA 180
DB
QY 181 LFLIDNRRRACHPCSPMKGSKRCWGESSEDCQSLTRTYCAGGCACRCKPLPTDCCHBOC 240
XX |
XX 181 LFLIDNRRRACHPCSPMKGSKRCWGESSEDCQSLTRTYCAGGCACRCKPLPTDCCHBOC 240
DB
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFESMPNPGRYTFGASCVTACP 300
XX |
XX 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFESMPNPGRYTFGASCVTACP 300
DB
QY 301 YNVLSTDVGSCTLVCPHLHNOEYTAEDGTQRCERCKSPCAR 340
XX |
XX 301 YNVLSTDVGSCTLVCPHLHNOEYTAEDGTQRCERCKSPCAR 340
DB
XX
XX RESULT 14
XX AA84780
XX ID AA84780 standard; Protein; 1255 AA.
XX
XX AC AA84780;
XX
XX DT 08-AUG-2000 (first entry)
XX
XX DE Amino acid sequence of the SPLICE erB-2 receptor protein.
XX
XX KW SPLICE erB-2 receptor protein; cell transformation disorder; cancer;
XX tumor cell proliferation; tissue degeneration; arthropathy;
XX bone resorption; inflammatory disease; degenerative disorder;
XX wound healing.
XX

```

```

OS Homo sapiens.
XX
XX MO200020579-A1.
XX
XX 13-APR-2000.
XX
XX PF 01-OCT-1999; 99WO-CA00912.
XX
XX PR 02-OCT-1998; 98US-0165192.
XX
XX (UYMC-) UNIV MCMMASTER.
XX
XX Muller WJ, Siegel PM;
XX
XX WPI: 2000-303768/26.
XX
XX N-PSDB: AAA14812.
XX
XX Nucleic acid encoding an erB 2 receptor protein designated SPLICE
XX erB-2, inhibitors of the protein are useful for treatment of cancer -
XX
XX Claim 3; Fig 2; 60pp; English.
XX
XX The present sequence represents a SPLICE erB-2 receptor protein. The
XX protein has an in-frame deletion of 16 amino acids, 2 of which are
XX conserved cysteine residues, compared to the unspliced protein. The
XX erB-2 polynucleotide is used to construct probes for detecting
XX disorders of cell transformation such as cancer. Antipodes to the
XX protein may be used to detect SPLICE erB-2 in a sample. Agents
XX (e.g. antisense oligonucleotides) which inhibit the expression of
XX SPLICE erB-2 are useful for reducing tumor cell proliferation and
XX for treating cancer. Substances which stimulate SPLICE erB-2 are useful
XX for treating conditions involving damaged cells including conditions
XX in which degeneration of tissue occurs, such as arthropathy, bone
XX resorption, inflammatory diseases, degenerative disorders of the
XX central nervous system and wound healing.
XX
XX Sequence 1255 AA;
XX
XX Query Match 81.1%; Score 340; DB 21; Length 1255;
XX Best Local Similarity 100.0%; Pred. No. 6.9e-302;
XX Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MELAAACRNGLLALLPPAASSTQVCTGDMKRLRPASETHLDMRLHXGCGVQGNL 60
XX |
XX 1 MELAAACRNGLLALLPPAASSTQVCTGDMKRLRPASETHLDMRLHXGCGVQGNL 60
DB
QY 61 ELTYLPTNASLFLDIQIEVQGYVLLAHNOVROVPLQRLRIYVGTQLFEDNVALAVLDNG 120
XX |
XX 61 ELTYLPTNASLFLDIQIEVQGYVLLAHNOVROVPLQRLRIYVGTQLFEDNVALAVLDNG 120
DB
QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQRPOLCYQDTILMKDIFHKNNOLA 180
XX |
XX 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQRPOLCYQDTILMKDIFHKNNOLA 180
DB
QY 181 LFLIDNRRRACHPCSPMKGSKRCWGESSEDCQSLTRTYCAGGCACRCKPLPTDCCHBOC 240
XX |
XX 181 LFLIDNRRRACHPCSPMKGSKRCWGESSEDCQSLTRTYCAGGCACRCKPLPTDCCHBOC 240
DB
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFESMPNPGRYTFGASCVTACP 300
XX |
XX 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDFESMPNPGRYTFGASCVTACP 300
DB
QY 301 YNVLSTDVGSCTLVCPHLHNOEYTAEDGTQRCERCKSPCAR 340
XX |
XX 301 YNVLSTDVGSCTLVCPHLHNOEYTAEDGTQRCERCKSPCAR 340
DB
XX
XX RESULT 15
XX AA92620
XX ID AA92620 standard; Protein; 1255 AA.
XX
XX AC AA92620;
XX

```

DT 10-AUG-2000 (first entry)
XX
DE Human heregulin 2 (Her2).
XX
XX Heregulin 2; Her2; vaccination; cytotoxic T-lymphocyte immunity;
KM self-protein; cancer; breast cancer; prostate cancer;
XX cell-associated peptide antigen; foreign epitope.
XX
OS Homo sapiens.
XX
FH Key
FH Domain
FT Location/Qualifiers
FT 1..173
FT /label= N-terminal
FT /note= "mature polypeptide"
FT 5..25
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 59..73
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 103..117
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 149..163
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 174..323
FT /label= Cysteine_rich_domain
FT 210..224
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 250..264
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 324..483
FT /label= Ligand_binding_domain
FT 325..339
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 369..383
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 465..479
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 484..623
FT /label= Cysteine_rich_domain
FT 579..593
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 624..654
FT /label= Transmembrane_domain
FT 632..652
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 653..667
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 655..1010
FT /label= Tyrosine_kinase_domain
FT 661..675
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 695..709
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 710..730
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 1011..1235
FT /label= C-terminal_domain
XX
PN WO200020027-A2.

XX 13-APR-2000.
PD
XX 05-OCT-1999; 99WO-DK00525.
XX
XX 05-OCT-1998; 98DK-0001261.
PR 20-OCT-1998; 98US-0105011.
XX
XX (MERI-) M & E BIOTECH AS.
XX
PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
DR WPI: 2000-349917/30.
DR N-PSDB; AAA09455.
XX
PT Inducing immune responses to weakly immunogenic, tumor associated
PT peptide antigens for the treatment of breast and prostate cancer
XX
PS Claim 62; Page 193-198; 220pp; English.
XX
CC This is the human heregulin 2 (Her2) sequence. Immunogenic analogues of
CC Her2 can be used in the claimed method as an autovaccine to induce a CTL
CC response. Subdominant CTL epitopes, antibody binding regions and
CC cysteine residues involved in disulfide bonds are preserved in the
CC immunogenized forms. Regions suitable for the insertion of foreign T
CC helper epitopes were identified (see features table). The method
CC is used for inducing immune responses against weakly immunogenic
CC cell-associated peptide antigens (PA) such as those associated with
CC cancers (self-proteins), e.g. human prostate specific membrane antigen
CC (PSM), heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b).
CC The method comprises effecting simultaneous presentation by antigen
CC producing cells (APCs) of the animals immune system of: (1) at least 1
CC CTL (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
CC B-cell group derived from the cell-associated PA; and (2) at least 1
CC first T helper cell group which is foreign to the animal. Analogues of
CC human PSM, human Her2 and human/murine FGF8b comprising a substantial
CC part of all known and predicted CTL and B-cell epitopes of the respective
CC PA and including at least one foreign T helper epitope are also claimed.
CC The method is used to treat prostate, prostate/breast or breast cancer
CC when the PA is human PSM, FGF8b and Her2, respectively.
XX
SQ Sequence 1255 AA;
Query Match 81.1%; Score 340; DB 21; Length 1255;
Best Local Similarity 100.0%; Pred. No. 6; 9e-302;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MELAALCRWGLLLALLPPGAATQVCTGDMKLRIPASETHLDMRLHYGCGVVGNTL 60
DB 1 MELAALCRWGLLLALLPPGAATQVCTGDMKLRIPASETHLDMRLHYGCGVVGNTL 60
QY 61 ELTYLPTNASTSFLDIOEVQGYVLIANHOVQVPLQRIARTYRGQLFEDNALAVLDNG 120
DB 61 ELTYLPTNASTSFLDIOEVQGYVLIANHOVQVPLQRIARTYRGQLFEDNALAVLDNG 120
QY 121 DPLNNTPTVTGASPGCLRELOLSLFEILKGYLIORNPOLCYODTILKKDIFHKNNOLA 180
DB 121 DPLNNTPTVTGASPGCLRELOLSLFEILKGYLIORNPOLCYODTILKKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMKGSRKMGSESDCSLRTVCAGGACARCKRPLPTDCCHQC 240
DB 181 LTLIDNRSRACHPCSPMKGSRKMGSESDCSLRTVCAGGACARCKRPLPTDCCHQC 240
QY 241 AAGCTGPKHSDDLACHFNHSGICELHCPALVTYNTDTRESMPNPEGRTFPASCYTACP 300
DB 241 AAGCTGPKHSDDLACHFNHSGICELHCPALVTYNTDTRESMPNPEGRTFPASCYTACP 300
QY 301 YNLTSDVGSCTLYVCPLHNOEYTAEDGTORCERCKSPCAR 340
DB 301 YNLTSDVGSCTLYVCPLHNOEYTAEDGTORCERCKSPCAR 340

Tue Mar 4 13:30:20 2003

us-09-234-208b-2.oli.rag

Page 11

Search completed: March 4, 2003, 12:44:57
Job time : 57.8474 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:43:27 ; Search time 47.1165 seconds
(without alignments)
854,910 Million cell updates/sec

Title: US-09-234-208b-2
Perfect score: 419
Sequence: 1 MELAALCRWGLLALLPQA.....VGRGPDPAHVAVNLRYEG 419

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	340	81.1	1255	1 A24571	protein-tyrosine k
2	67	16.0	1254	2 I48161	p-185 precursor
3	52	12.4	1260	1 TVRTNU	protein-tyrosine k
4	9	2.1	230	2 B82252	knife-related prote
5	9	2.1	527	2 A42032	epidermal growth f
6	9	2.1	644	2 A36325	epidermal growth f
7	9	2.1	1210	1 G0HUE	epidermal growth f
8	9	2.1	1210	2 A53183	epidermal growth f
9	9	2.1	1223	1 TVCHLV	epidermal growth f
10	8	1.9	277	2 JC5284	carboxyl reductase
11	8	1.9	341	2 B83298	conserved hypotet
12	8	1.9	348	2 D95067	phenylalanine-trna
13	8	1.9	375	2 C97935	phenylalanine-trna
14	8	1.9	463	2 H70922	hypothetical prote
15	8	1.9	520	2 C70311	hypothetical prote
16	8	1.9	583	2 D90052	hypothetical prote
17	8	1.9	725	2 IJMSNG	hypothetical prote
18	8	1.9	858	1 IJRTNC	neural cell adhesi
19	8	1.9	865	2 A85032	neural cell adhesi
20	8	1.9	879	2 S43910	hypothetical prote
21	8	1.9	1115	1 IJMSNL	chloroplast outer
22	8	1.9	1503	2 T01098	neural cell adhesi
23	7	1.7	62	2 F90177	chloroplast outer
24	7	1.7	104	2 G69482	hypothetical prote
25	7	1.7	111	2 C35826	hypothetical prote
26	7	1.7	121	2 AH1848	hypothetical 13k p
27	7	1.7	173	2 S10199	hypothetical prote
28	7	1.7	189	2 A70186	NADH2 dehydrogenas
29	7	1.7	189	2 S04670	neutrophil activat
					hypothetical prote

30	7	1.7	207	2 A69941	capsular polysacch
31	7	1.7	209	2 T00733	hypothetical prote
32	7	1.7	220	2 S05595	trypsin inhibitor
33	7	1.7	220	2 I50588	fibroblast growth
34	7	1.7	221	2 UC7587	stromal cell-deriv
35	7	1.7	223	2 H83462	heme exporter prot
36	7	1.7	224	1 A33861	trans-activating t
37	7	1.7	248	2 C83431	type III export pr
38	7	1.7	255	2 AH2460	glycosyltransferas
39	7	1.7	259	2 D75275	endonuclease III -
40	7	1.7	268	2 C70620	hypothetical prote
41	7	1.7	268	2 C90709	thiol disulfide in
42	7	1.7	268	2 G85559	thiol disulfide in
43	7	1.7	268	2 B64794	hypothetical prote
44	7	1.7	269	1 D43706	serine O-acetyltra
45	7	1.7	271	2 A45606	DNA-binding protei

ALIGNMENTS

RESULT 1
A24571
protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human
N:Alternate names: c-erb-B-2 protein precursor; Kinase-related transforming protein e
C:Species: Homo sapiens (man)
C>Date: 25-Oct-1987 #sequence_revision 06-Dec-1996 #text_change 11-Jun-1999
C:Accession: A24571, A25491, A44188, B44188, I59509, I57622
R:Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T
Nature 319, 230-234, 1986
A>Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growt
A:Reference number: A24571; MUID:86118663; PMID:3003577
A:Accession: A24571
A:Molecule type: mRNA
A:Residues: 1-1255 <YAM>
A:Cross-references: GB:X03363; NID:g31197; PIDN:CAA27060.1; PID:g31198
R:Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.
Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985
A>Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epid
A:Reference number: A25491; MUID:86016729; PMID:2995967
A:Accession: A25491
A:Molecule type: DNA
A:Residues: 737-1031 <SEM>
A:Cross-references: GB:M11767; NID:g182163; PIDN:AAA35808.1; PID:g553282
R:Cousens, L.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg
Science 230, 1132-1139, 1985
A>Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chro
A:Reference number: A44188; MUID:86070181; PMID:2999974
A:Accession: A44188
A:Molecule type: DNA
A:Residues: 740-910 <COU1>
A:Cross-references: GB:M12036; NID:g183988; PIDN:AAA35978.1; PID:g183989
A:Accession: B44188
A:Molecule type: mRNA
A:Residues: 1-517, 'RALF', 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>
A:Cross-references: GB:M1730; NID:g183986
R:King, C.R.; Kraus, M.H.; Aaronson, S.A.
Science 229, 974-976, 1985
A>Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.
A:Reference number: I59509; MUID:85272597; PMID:2992089
A:Accession: I59509
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 832-909 <REX>
A:Cross-references: GB:L29395; NID:g459807; PIDN:AAA35809.1; PID:g459808
R:Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D.
Mol. Cell. Biol. 7, 2597-2601, 1987
A>Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcrip
A:Reference number: I57622; MUID:87286688; PMID:3039551
A:Accession: I57622
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-191 <TAL>

A:Cross-references: GB:M16792; NID:g183983; PIDN:AA58637.1; PID:g553332
C:Comment: Amplification and overexpression of this erbB-related gene occurs in about 30
C:Genetics:
A:Gene: GDB:ERBB2; NGU: NEU: HER-2
A:Cross-references: GDB:120613; OMIM:164870
A:Map position: 17q21.1-17q21.1
A:Introns: 25/1; 75/3; 147/1; 883/3
A>Note: the 1st of introns is incomplete
C:Function:
A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphatase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-125/Product: protein-tyrosine kinase erbB2 #status predicted <MAT>
F:22-653/Domain: extracellular #status predicted <EXT>
F:70-304/Domain: EGF receptor extracellular domain repeat <EE1>
F:395-605/Domain: EGF receptor extracellular domain repeat <EE2>
F:654-675/Domain: transmembrane #status predicted <TM>
F:676-1255/Domain: intracellular #status predicted <INT>
F:718-983/Domain: protein kinase homology <KIN>
F:726-734/Region: protein kinase ATP-binding motif
F:68-124,187,259,530,571,629/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
F:753/Active site: Lys #status predicted
F:1139,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 81.1%; Score 340; DB 1; Length 1255;

Best Local Similarity 100.0%; Pred. No. 0; Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 METALALCRLGGLLALPPGAASVTCGTDMKRLPASPEHLMRLHYOGCGVGNL 60
Db 1 METALALCRLGGLLALPPGAASVTCGTDMKRLPASPEHLMRLHYOGCGVGNL 60
QY 61 ELTYLPNALSFLQDIOEVGYLIHNOYROYPLDRLKIVKSTOLFEDNYALAVDNG 120
Db 61 ELTYLPNALSFLQDIOEVGYLIHNOYROYPLDRLKIVKSTOLFEDNYALAVDNG 120
QY 121 DPLNNTPTVVGASPGSGREVLRLSLTEILKGVLIQRNPOLCQDITLMDIFHKNNOLA 180
Db 121 DPLNNTPTVVGASPGSGREVLRLSLTEILKGVLIQRNPOLCQDITLMDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMGKSGRCWGESSEDDCSLRTVACGAGCARCKGLPTDCHEQC 240
Db 181 LTLIDNRSRACHPCSPMGKSGRCWGESSEDDCSLRTVACGAGCARCKGLPTDCHEQC 240
QY 241 AAGCTGKHSDDCLACLFHNSGICELHCPALVYNTDTFESMPREGRYTFGASCVTACP 300
Db 241 AAGCTGKHSDDCLACLFHNSGICELHCPALVYNTDTFESMPREGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPAHNOEVTAEADTGRCCKSPCAR 340
Db 301 YNYLSTDVGSCTLVCPAHNOEVTAEADTGRCCKSPCAR 340

RESULT 2

148161

P:185 precursor - golden hamster

C:Species: Mesocricetus auratus (golden hamster)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999

C:Accession: 148161

R:Nakamura, T.; Ushijima, T.; Ishizaka, Y.; Nagao, M.; Arai, M.; Yamazaki, Y.; Ishikawa, Gene 140, 251-255, 1994

A>Title: Cloning and activation of the Syrian hamster neu proto-oncogene.

A:Reference number: 148161; MIMD:94193007; PMID:7908275

A:Accession: 148161

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1254 <RES>

C:Cross-references: GB:D16295; NID:g493236; PIDN:BA03801.1; PID:g747595

C:Genetics:
A:Gene: neu

C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: ATP
F:718-983/Domain: protein kinase homology <KIN>
F:726-734/Region: protein kinase ATP-binding motif

Query Match 16.0%; Score 67; DB 2; Length 1254;

Best Local Similarity 100.0%; Pred. No. 7,4e-58; Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 LPTDCHGQCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPREGRYT 290
Db 231 LPTDCHGQCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPREGRYT 290

QY 291 FGASCVT 297

Db 291 FGASCVT 297

RESULT 3

TVRTNU

C:Species: Rattus norvegicus (Norway rat)

C>Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 11-Jun-1999

C:Accession: A24562; A61204

R:Barthmann, C.I.; Hung, M.C.; Weinberg, R.A.

Nature 319, 226-230, 1986

A>Title: The new oncogene encodes an epidermal growth factor receptor-related protein

A:Reference number: A24562; MIMD:86118662; PMID:3945311

A:Accession: A24562

A:Molecule type: mRNA

A:Residues: 1-1260 <BAR>

A:Cross-references: EMBL:X03362; NID:g56745; PIDN:CAA27059.1; PID:g56746

R:Masui, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohe Carcinogenesis 12, 1975-1978, 1991

A>Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals a 2-thiazolyl[1]formamide or N-methyl-N-nitrosurea.

A:Reference number: A61204; MIMD:92035293; PMID:1682063

A:Accession: A61204

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 637-663, 'V', 665-702 <MAS>

A>Note: authors translated the codon GCA for residue 25 as Val

C:Genetics:

C:Superfamily: epidermal growth factor receptor; protein kinase homology

C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phos

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-1260/Product: protein-tyrosine kinase neu #status predicted <MAT>

F:658-680/Domain: transmembrane #status predicted <TM>

F:723-988/Domain: protein kinase ATP-binding motif

F:731-739/Region: protein kinase ATP-binding motif

F:71,191,263,535,576,634/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:691/Binding site: phosphate (Thr) (covalent) #status predicted

F:758/Active site: Lys #status predicted

F:862,1227,1233/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 12.4%; Score 52; DB 1; Length 1260;

Best Local Similarity 100.0%; Pred. No. 6,2e-43; Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 LPTDCHGQCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESM 282
Db 235 LPTDCHGQCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESM 286

RESULT 4

B82252

R:ite-related protein VC01012 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C:Species: Vibrio cholerae

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: B82252

R:Heideberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R. Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoli, I.; Sellers

L.R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
 A:Reference number: A82035; MUID:20406833; PMID:10952301
 A:Accession: B82252

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-230 <HEI>
 A:Cross-references: GB:AE004183; GB:AE003852; NID:g9655473; PIDN:AAF94173.1; GSPDB:GN001
 A:Experimental source: serogroup O1, strain N16961, biotype El Tor
 C:Genetics:
 A:Gene: VC1012
 A:Map position: 1
 C:Superfamily: conserved hypothetical protein H11688

Query Match 2.1%; Score 9; DB 2; Length 230;
 Best Local Similarity 100.0%; Pred. No. 0.82;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 LVALPPGA 20
 |||||
 Db 184 LVALPPGA 192

RESULT 5
 A42032
 epidermal growth factor receptor - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 18-Jun-1999
 C:Accession: A42032
 R:Flickinger, T.W.; Maithe, N.J.; Kung, H.J.
 Mol. Cell. Biol. 12, 883-893, 1992
 A:Title: An alternatively processed mRNA from the avian c-erbB gene encodes a soluble, t
 A:Reference number: A42032; MUID:92123214; PMID:1732751
 A:Accession: A42032

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-527 <FLI>
 A:Cross-references: GB:M77637; NID:g211737; PIDN:AAA48759.1; PID:g211738
 A:Experimental source: liver
 A:Note: sequence extracted from NCBI backbone (NCBIN:76692, NCBI:76693)
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; growth factor receptor

Query Match 2.1%; Score 9; DB 2; Length 527;
 Best Local Similarity 100.0%; Pred. No. 1.7;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 239 OCAAGCTGP 247
 |||||
 Db 245 OCAAGCTGP 253

RESULT 6
 A36325
 epidermal growth factor receptor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 25-Jan-1991 #sequence_revision 25-Jan-1991 #text_change 10-Oct-1997
 C:Accession: A36325
 R:Petich, L.A.; Harris, J.; Raymond, V.W.; Blasband, A.; Lee, D.C.; Earp, H.S.
 Mol. Cell. Biol. 10, 2973-2982, 1990
 A:Title: A truncated, secreted form of the epidermal growth factor receptor is encoded b
 A:Reference number: A36325; MUID:90258888; PMID:2342466
 A:Accession: A36325

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-644 <PER>
 A:Cross-references: GB:M37394
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: alternative splicing; ATP; growth factor receptor

Query Match 2.1%; Score 9; DB 2; Length 644;
 Best Local Similarity 100.0%; Pred. No. 2;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 239 OCAAGCTGP 247
 |||||
 Db 235 OCAAGCTGP 243

RESULT 7
 G0HUE
 epidermal growth factor receptor precursor - human
 N:Contains: protein-tyrosine kinase (EC 2.7.1.112) erbB
 C:Species: Homo sapiens (man)
 C:Date: 15-Nov-1984 #sequence_revision 27-Nov-1985 #text_change 11-Jun-1999
 C:Accession: A00641; A25772; S30024; A38672; A00642; A43615; A23062; A05281; A60143;
 R:Ulrich, A.; Coussens, L.; Hayflick, J.S.; Dull, T.J.; Gray, A.; Tam, A.W.; Lee, J.
 rg, P.H.
 Nature 309, 418-425, 1984
 A:Title: Human epidermal growth factor receptor cDNA sequence and aberrant expression
 A:Reference number: A00641; MUID:84219729; PMID:6328312
 A:Accession: A00641

A:Molecule type: mRNA
 A:Residues: 1-1210 <UUL>
 A:Cross-references: EMBL:X00588; NID:g31113; PIDN:CAA25240.1; PID:g757924
 A:Note: the authors translated the codon AAG for residue 540 as Asn
 R: Ishii, S.; Xu, Y.; Stratton, R.H.; Roe, B.A.; Merlino, G.T.; Pastan, I.
 Proc. Natl. Acad. Sci. U.S.A. 82, 4920-4924, 1985
 A:Title: Characterization and sequence of the promoter region of the human epidermal
 A:Reference number: A25772; MUID:85270438; PMID:2991899
 A:Accession: A25772

A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-29 <ISH>
 A:Cross-references: GB:M11234; NID:g181981; PIDN:AAA52370.1; PID:g553272
 R:Haley, J.; Whittle, N.; Bennett, P.; Kinchington, D.; Ulrich, A.; Waterfield, M.
 Oncogene Res. 1, 375-396, 1987
 A:Title: The human EGF receptor gene: structure of the 110 kb locus and identification
 A:Reference number: S30024; MUID:88217333; PMID:3329716
 A:Accession: S30024

A:Molecule type: DNA
 A:Residues: 1-29 <HA2>
 A:Cross-references: EMBL:X06370; NID:g31118; PIDN:CAA29668.1; PID:g31119
 R:Haley, J.D.; Waterfield, M.D.
 J. Biol. Chem. 266, 1746-1753, 1991
 A:Title: Contributory effects of de Novo transcription and premature transcript termi
 A:Reference number: A38672; MUID:91107677; PMID:1988448
 A:Accession: A38672

A:Molecule type: DNA
 A:Residues: 1-29 <HAL>
 A:Cross-references: GB:M38425; NID:g181977; PIDN:AAA63171.1; PID:g553271
 R:Xu, Y.; Ishii, S.; Clark, A.J.L.; Sullivan, M.; Wilson, R.K.; Ma, D.P.; Roe, B.A.;
 Nature 309, 806-810, 1984
 A:Title: Human epidermal growth factor receptor cDNA is homologous to a variety of RN
 A:Reference number: A00642; MUID:84245835; PMID:6330563
 A:Accession: A00642

A:Molecule type: mRNA
 A:Residues: 713-964 <LIN>
 A:Experimental source: epidermoid carcinoma cell line A431
 R:Stimmen, F.A.; Gope, M.L.; Schulz, T.Z.; Wright, D.A.; Carpenter, G.; O'Malley, B.W.
 Biochem. Biophys. Res. Commun. 124, 125-132, 1984
 A:Reference number: A23062; MUID:85046483; PMID:6093780
 A:Accession: A23062
 A:Molecule type: mRNA
 A:Residues: 1028-1210 <SIM>

R: Weber, W.; Gill, G.N.; Speiss, J.
 Science 224, 294-297, 1984
 A:Reference number: A05281; MUID:84172183; PMID:6324343
 A:Accession: A05281
 A:Molecule type: protein
 A:Residues: 25-30, 'S', '32-51, 454-467 <WEB>
 R: Russo, M.W.; Lukas, T.J.; Cohen, S.; Staros, J.V.
 J. Biol. Chem. 260, 5205-5208, 1985
 A:Title: Identification of residues in the nucleotide binding site of the epidermal growth factor receptor
 A:Reference number: A60143; MUID:85182650; PMID:2985580
 A:Accession: A60143
 A:Molecule type: protein
 A:Residues: 740-744, 'X', '746-747 <RUS>
 R: Mroczkowski, B.; Mosig, G.; Cohen, S.
 Nature 309, 270-273, 1984
 A:Title: ATP-stimulated interaction between epidermal growth factor receptor and superoxide dismutase
 A:Reference number: A58023; MUID:84191554; PMID:6325948
 A:Contents: annotation; receptor activity
 A:Note: The EGF receptor (and other tyrosine kinases) can nick double-stranded DNA
 R: Chen, W.S.; Lazar, C.S.; Lund, K.A.; Welsh, J.B.; Chang, C.P.; Walton, G.M.; Der, C.J.
 Cell 59, 33-43, 1989
 A:Title: Functional independence of the epidermal growth factor receptor from a domain involved in ligand binding
 A:Reference number: A53331; MUID:90003233; PMID:2790960
 A:Contents: annotation; internalization signal
 C:Comment: Binding of EGF to the receptor leads to internalization of the EGF-receptor complex
 A:Gene: GDB:EGFR
 A:Cross-references: GDB:120610; OMIM:131550
 A:Map position: 7p12.3-7p21.1
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphotyrosine; tyrosine phosphorylation
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-1210/Product: EGF receptor #status predicted <MAT>
 F:25-645/Domain: extracellular #status predicted <EXT>
 F:75-300/Domain: EGF receptor extracellular domain repeat <EET>
 F:390-600/Domain: EGF receptor extracellular domain repeat <EE2>
 F:646-668/Domain: transmembrane #status predicted <TM>
 F:669-1210/Domain: intracellular #status predicted <INT>
 F:710-975/Domain: protein kinase homology <KIN>
 F:718-726/Region: protein kinase ATP-binding motif
 F:998-1046/Region: coated-pit mediated internalization signal
 F:1047-1210/Region: inhibitory
 F:128,175,352,413,444,528,603/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:745/Active site: Lys #status experimental
 Query Match 2.1%; Score 9; DB 1; Length 1210;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 239 OCAAGCTGP 247
 DB 235 OCAAGCTGP 243

RESULT 8
 A53183
 epidermal growth factor receptor precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 18-Jun-1999
 C:Accession: A53183; A43818; S24942; A28941; S45325; I49643
 R: Luettich, N.C.; Phillips, H.K.; Qiu, T.H.; Copeland, N.G.; Earp, H.S.; Jenkins, N.A.; Genes Dev. 8, 399-413, 1994
 A:Title: The mouse waved-2 phenotype results from a point mutation in the EGF receptor tyrosine kinase domain
 A:Reference number: A53183; MUID:94170986; PMID:8125255
 A:Accession: A53183
 A:Molecule type: mRNA
 A:Residues: 1-1210 <LUE>
 A:Cross-references: GB:U03425
 R: Avioli, A.; Lax, I.; Ullrich, A.; Schlessinger, J.; Givol, D.; Morse, B.
 Oncogene 6, 673-676, 1991
 A:Title: Comparison of EGF receptor sequences as a guide to study the ligand binding site
 A:Reference number: A43818; MUID:91232866; PMID:2030916
 A:Accession: A43818

A:Molecule type: mRNA
 A:Residues: 1-714 <AVI>
 A:Cross-references: GB:X59698
 R: Eisinger, D.P.; Serrero, G.
 submitted to the EMBL Data Library, June 1992
 A:Reference number: S24942
 A:Accession: S24942
 A:Molecule type: mRNA
 A:Residues: 969-971, 'K', '973-1115, 'D' <EIS>
 A:Cross-references: EMBL:Z12608
 R: Heisermann, G.J.; Gill, G.N.
 J. Biol. Chem. 263, 13152-13158, 1988
 A:Title: Epidermal growth factor receptor threonine and serine residues phosphorylate the kinase domain
 A:Reference number: A28941; MUID:88330814; PMID:3138233
 A:Accession: A28941
 A:Molecule type: protein
 A:Residues: 689-694, 'X', '696-704, 'L', '706-707, 989-992, 'XX', '995-996, 'X', '998-1000, 1002-1003
 R: Hibbs, M.L.; Dunn, A.R.; Alexander, W.S.
 submitted to the EMBL Data Library, April 1994
 A:Description: The complete cDNA sequence of the Mouse Epidermal Growth Factor Receptor
 A:Reference number: S45325
 A:Accession: S45325
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-971, 'K', '973-1210 <VER>
 A:Cross-references: EMBL:X78987; NID:q488830; PIDN:CAA55587.1; PID:q488831
 R: Paria, B.C.; Das, S.K.; Andrews, G.K.; Dey, S.K.
 Proc. Natl. Acad. Sci. U.S.A. 90, 55-59, 1993
 A:Title: Expression of the epidermal growth factor receptor gene is regulated in mouse embryonic tissues
 A:Reference number: I49643; MUID:93126380; PMID:7678348
 A:Accession: I49643
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 12-20, 22-132 <RES>
 A:Cross-references: GB:L06864; NID:q193001; PIDN:AAA53029.1; PID:q567201
 A:Gene: EGFR
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; growth factor receptor; kinase-related transforming protein; phosphotyrosine; tyrosine phosphorylation
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:648-670/Domain: transmembrane #status predicted <TM>
 F:712-977/Domain: protein kinase homology <KIN>
 F:720-728/Region: protein kinase ATP-binding motif
 F:680,695/Binding site: phosphate (Thr) (covalent) #status experimental
 F:697,1070,1071/Binding site: phosphate (Ser) (covalent) #status experimental
 F:993/Binding site: (or 997) phosphate (Ser) (covalent) #status experimental
 F:1028/Binding site: (or 1030 or 1032) phosphate (Ser) (covalent) #status experimental
 F:1157/Binding site: phosphate (Tyr) (covalent) #status experimental
 Query Match 2.1%; Score 9; DB 2; Length 1210;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 239 OCAAGCTGP 247
 DB 235 OCAAGCTGP 243

RESULT 9
 TVCHLV
 epidermal growth factor receptor precursor - chicken
 N:Contains: protein-tyrosine kinase (EC 2.7.1.112) erbB
 C:Species: Gallus gallus (chicken)
 C:Date: 28-Feb-1996 #sequence_revision 05-May-1995 #text_change 04-Feb-2000
 C:Accession: A27720; A00643
 R: Lax, I.; Johnson, A.; Howk, R.; Sap, J.; Bellot, F.; Winkler, M.; Ullrich, A.; Venn
 Mol. Cell. Biol. 8, 1970-1978, 1988
 A:Title: Chicken epidermal growth factor (EGF) receptor: cDNA cloning, expression in COS cells
 A:Reference number: A27720; MUID:88261272; PMID:3260329
 A:Accession: A27720
 A:Molecule type: mRNA
 A:Residues: 1-1223 <LAX>
 A:Cross-references: GB:M20386

R.Nilsen, T.W.; Maroney, P.A.; Goodwin, R.G.; Rottman, F.M.; Ciltenden, L.B.; Raines, M.
Cell 41, 719-726, 1985
A:Title: c-erbB activation in ALV-induced erythroblastosis: novel RNA processing and pro
A:Reference number: A00643; MUID:85228222; PMID:2988784
A:Accession: A00643
A:Molecule type: mRNA
A:Residues: 585-1223 <NIL>
A:Cross-references: GB:M10066
C:Genetics:
A:Gene: erbB
C:superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth factor
specific protein kinase
F:1-30/Domain: signal sequence #status predicted <SIG>
F:1-1223/Product: epidermal growth factor receptor #status predicted <MAP>
F:31-654/Domain: extracellular #status predicted <EX>
F:81-307/Domain: EGF receptor extracellular domain repeat <E1>
F:397-610/Domain: EGF receptor extracellular domain repeat <E2>
F:655-677/Domain: transmembrane #status predicted <TM>
F:678-1223/Domain: intracellular #status predicted <INT>
F:719-984/Domain: protein kinase homology <KIN>
F:727-735/Region: protein kinase ATP-binding motif
F:136,202,280,361,370,422,575,580,615,635/Binding site: carbohydrate (Thr) (covalent) #s
F:192,650/Binding site: carbohydrate (Ser) (covalent) #status predicted
F:687/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
F:754/Active site: Lys #status predicted
F:1100,1183,1208/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #stat

Query Match 2.1%; Score 9; DB 1; Length 1223;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 239 QCAAGCTGP 247
DB 242 QCAAGCTGP 250

RESULT 10
JC5284
C:Species: Rattus norvegicus (Norway rat)
C:Date: 16-Apr-1997 #sequence_revision 18-Jul-1997 #text_change 03-Jun-2002
C:Accession: S68982; PC2234; JC5284; PC2159; S52349
R:Wernuth, B.; Maeder-Helmann, G.; Ernst, E.
Eur. J. Biochem. 228, 473-479, 1995
A:Title: Cloning and expression of carbonyl reductase from rat testis.
A:Reference number: S68982; MUID:95220378; PMID:7705364
A:Accession: S68982
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-277 <WER>
A:Cross-references: EMBL:X84349; NID:g666086; PIDN:CA59088.1; PID:g666087
R:Toft, E.; Soederstrom, M.; Ahlberg, M.B.; DePierre, J.W.
Biochem. Biophys. Res. Commun. 201, 149-154, 1994
A:Title: A novel 34kDa glutathione-binding protein in mature rat ovary.
A:Reference number: PC2159; MUID:94256971; PMID:8198567
A:Accession: PC2234
A:Molecule type: protein
A:Residues: 104-121,'X','123-134','D','136-137','H' <TOF>
A:Experimental source: ovary
R:Aoki, H.; Okada, T.; Mizutani, T.; Numata, Y.; Minegishi, T.; Miyamoto, K.
Biochem. Biophys. Res. Commun. 230, 518-523, 1997
A:Title: Identification of two closely related genes, inducible and noninducible carbonyl
A:Reference number: JC5284; MUID:97167735; PMID:9015353
A:Contents: ovary
A:Accession: JC5284
A:Molecule type: mRNA
A:Residues: 1-140,'GM','143','R','145-235','T','237-238','E','240-277' <AOK>
A:Cross-references: DBJ:B89069; NID:g1906811; PIDN:BA19007.1; PID:g1906812
C:Comment: This enzyme is a cytosolic, monomeric oxidoreductase that catalyzes the NADPH
C:Genetics:
A:Gene: ICR
C:superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

C:Keywords: cytosol; monomer; NADP; oxidoreductase
F:6-185/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 1.9%; Score 8; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 LALLPGA 20
DB 254 LALLPGA 261

RESULT 11
B83298
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B83298
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pa
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: B83298
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-341 <STO>
A:Cross-references: GB:AE004705; GB:AE004091; NID:g9948851; PIDN:AAG06159.1; GSPDB:GN
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: PA2771

Query Match 1.9%; Score 8; DB 2; Length 341;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 140 LQRLSLE 147
DB 158 LQRLSLE 165

RESULT 12
D95067
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C:Accession: D95067
R:Petelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Redune, D.; Holtzapp
son, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: D95067
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-348 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK74733.1; PID:g14972054; GSPDB:GN00164; TIGR
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP0579
C:superfamily: phenylalanine-tRNA ligase alpha chain

Query Match 1.9%; Score 8; DB 2; Length 348;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 144 SLTEILKG 151
DB 43 SLTEILKG 50

RESULT 13

C97935
phenylalanine-tRNA ligase (EC 6.1.1.20) [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C>Date: 22-Oct-2001 #sequence-revision 22-Oct-2001 #text_change 03-Jun-2002
C:Accession: C97935
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; E
y, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A>Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: C97935
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-375 <KUR>
A:Cross-references: GB:AE007317; PIDN:AAK99311.1; PID:g15458079; GSPDB:GN00174
C:Genetics:
A:Gene: phs
C:Superfamily: phenylalanine-tRNA ligase alpha chain
C:Keywords: .ligase

Query Match 1.9%; Score 8; DB 2; Length 375;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 144 SLTEILKG 151
Db 70 SLTEILKG 77

RESULT 14

H70922
hypothetical protein RV3130C - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence-revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: H70922
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: H70922
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-463 <COL>
A:Cross-references: GB:Z95150; GB:AL123456; NID:g3250708; PIDN:CAE08399.1; PID:el301034;
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV3130C

Query Match 1.9%; Score 8; DB 2; Length 463;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 95 PIQRLRIV 102
Db 315 PIQRLRIV 322

RESULT 15

C70311
hypothetical protein ag_116 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C>Date: 08-May-1998 #sequence-revision 08-May-1998 #text_change 09-Jun-2000
C:Accession: C70311
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O
V.
Nature 392, 353-358, 1998

A>Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A:Reference number: A70300; MUID:98196666; PMID:9537320

A:Accession: C70311

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-520 <AOE>

A:Cross-references: GB:AE000674; NID:g2982850; PIDN:AAE06482.1; PID:g2982861; GB:AE00

A:Experimental source: strain VF5

C:Genetics:

A:Gene: ag_116

C:Superfamily: Aquifex aeolicus hypothetical protein ag_116

Query Match 1.9%; Score 8; DB 2; Length 520;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 387 PLSPTSVP 394
Db 2 PLSPTSVP 9

Search completed: March 4, 2003, 12:47:43
Job time : 49.1165 secs.

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:37:16 ; Search time 28.6064 Seconds
(without alignments)
607.506 Million cell updates/sec

Title: US-09-234-208B-2

Perfect score: 419
Sequence: 1 MELALCRMGDLALLPPGA.....VGRGPPDAHVAVNLRYEG 419

Scoring table: OLIGO
Gapop 60.0 , Gapept 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	340	81.1	1255	1	ERR2_HUMAN
2	67	16.0	1254	1	ERR2_MESAU
3	52	12.4	1257	1	ERR2_RAT
4	9	2.1	230	1	RNFE_VIBCH
5	9	2.1	703	1	EGFR_CHICK
6	9	2.1	1210	1	EGFR_HUMAN
7	9	2.1	1210	1	EGFR_MOUSE
8	8	1.9	276	1	DHCA_RAT
9	8	1.9	463	1	YV30_MYCTU
10	8	1.9	725	1	NCA2_MOUSE
11	8	1.9	858	1	NCA1_RAT
12	8	1.9	1115	1	SECE_MOUSE
13	7	1.7	62	1	SECE_SULSO
14	7	1.7	104	1	Y164_ARCFU
15	7	1.7	111	1	NAD2_HUMAN
16	7	1.7	114	1	GON2_TUPGB
17	7	1.7	173	1	N16M_CHICK
18	7	1.7	189	1	YAT5_RHOBL
19	7	1.7	192	1	UREE_ALCEU
20	7	1.7	207	1	YPOP_BACSU
21	7	1.7	220	1	GF33_CHICK
22	7	1.7	221	1	SDFL_HUMAN
23	7	1.7	224	1	TCMD_SALTY
24	7	1.7	228	1	YU02_HUMAN
25	7	1.7	248	1	DSBG_ECOS7
26	7	1.7	248	1	DSBG_ECOLI
27	7	1.7	269	1	N1P7_AZOCH
28	7	1.7	271	1	HMG1_TRYBR
29	7	1.7	287	1	YCKE_BACSU
30	7	1.7	306	1	SRP1_SYNP7
31	7	1.7	307	1	MNP1_MYCLE
32	7	1.7	313	1	ENV1_MOUSE
33	7	1.7	334	1	G3P1_BACSU

34	7	1.7	341	1	XERD_SELNU
35	7	1.7	357	1	MLTC_HAEIN
36	7	1.7	365	1	Y271_YEAST
37	7	1.7	379	1	IRKA_RAT
38	7	1.7	380	1	YMH7_CABEL
39	7	1.7	383	1	PODO_HUMAN
40	7	1.7	396	1	YD18_YEAST
41	7	1.7	399	1	BRS3_SHEEP
42	7	1.7	400	1	RENT_SHEEP
43	7	1.7	423	1	ACT_HUMAN
44	7	1.7	443	1	CG1H_XANNA
45	7	1.7	443	1	SOX3_HUMAN

ALIGNMENTS

RESULT 1	ID	ERR2_HUMAN	STANDARD:	PRT: 1255 AA.
AC	P04626:			
DT	13-AUG-1987	(Rel. 05, Created)		
DT	13-AUG-1987	(Rel. 05, Last sequence update)		
DT	15-JUN-2002	(Rel. 41, Last annotation update)		
DE	Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)			
DE	(p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Tyrosine kinase-type cell surface receptor HER2) (MLN 19).			
GN	ERBB2 OR HER2 OR NGL OR NEU.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86118663; PubMed=3003577;			
RA	Yamamoto T., Ikawa S., Akiyama T., Semba K., Nomura N., Miyajima N.,			
RA	Saito T., Toyoshima K.;			
RT	"Similarity of protein encoded by the human c-erbB-2 gene to			
RT	epidermal growth factor receptor.";			
RL	Nature 319:230-234(1986).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86016729; PubMed=2999974;			
RA	Coussens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A.,			
RA	McGrath J., Seeburg P.H., Libermann T.A., Schlessinger J.,			
RA	Francke U., Levinson A., Ullrich A.;			
RT	"Tyrosine kinase receptor with extensive homology to EGF receptor			
RT	shares chromosomal location with neu oncogene.";			
RL	Science 230:1132-1139(1985).			
RN	[3]			
RP	SEQUENCE OF 737-1031 FROM N.A.			
RX	MEDLINE=86016729; PubMed=2999974;			
RA	Sema K., Kamata N., Toyoshima K., Yamamoto T.;			
RT	"A v-erbB-related protooncogene, c-erbB-2, is distinct from the			
RT	c-erbB-1/epidermal growth factor-receptor gene and is amplified in a			
RL	human salivary gland adenocarcinoma.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).			
RN	[4]			
RP	VARIANTS VAL-654 AND VAL-655.			
RX	MEDLINE=93194196; PubMed=8095488;			
RA	Ehsani A., Low J., Wallace R.B., Wu A.M.;			
RT	"Characterization of a new allele of the human ERBB2 gene by allele-			
RT	specific competition hybridization.";			
RL	Genomics 15:426-429(1993).			
CC	-i- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,			
CC	ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A			
CC	POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-			
CC	ALPHA AND AMHIREGULIN.			
CC	-i- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein			
CC	tyrosine phosphate.			
CC	-i- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS			
CC	(POTENTIAL).			
CC	-i- SUBCELLULAR LOCATION: Type I membrane protein.			

```

CC -i- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
CC RESIDUES (BY SIMILARITY).
CC -i- POLYMORPHISM: THERE ARE FOUR ALLELES DUE TO THE VARIATIONS IN
CC POSITIONS 654 AND 655. ALLELE B1 (654-ILE-ILE-655) HAS A FREQUENCY
CC OF 0.782; ALLELE B2 (654-ILE-ILE-655) HAS A FREQUENCY OF 0.206;
CC ALLELE B3 (654-VAL-VAL-655) HAS A FREQUENCY OF 0.012.
CC -i- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M11767; AAA35808.1; -
CC EMBL: M11761; AAA35808.1; JOINED.
CC EMBL: M11762; AAA35808.1; JOINED.
CC EMBL: M11763; AAA35808.1; JOINED.
CC EMBL: M11764; AAA35808.1; JOINED.
CC EMBL: M11765; AAA35808.1; JOINED.
CC EMBL: M11766; AAA35808.1; JOINED.
CC EMBL: M11730; AAA75493.1; -
CC EMBL: M12036; AAA35978.1; -
CC EMBL: X03363; CAA27060.1; -
CC PIR: A25491; A25491.
CC PIR: A24571; A24571.
CC HSP: P11362; 1FGK.
CC Genew: HGNC:3430; ERBB2.
CC MIM: 164870; -
CC InterPro: IPR000494; EGFR_L_domain.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR002174; Furin-like.
CC InterPro: IPR001245; Tyr_pkinase.
CC InterPro: IPR004019; YLP_motif.
CC Pfam: PF00069; pkinase; 1.
CC Pfam: PF00757; Furin-like; 1.
CC Pfam: PF01030; Recep_L_domain; 2.
CC Pfam: PF02757; YLP; 2.
CC Prodom: PD000001; Euk_pkinase; 1.
CC SMART: SM00261; FU; 3.
CC SMART: SM00219; TYRK; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
CC Transmembrane: Glycoprotein: Multigene family: Receptor; Signal:
CC Transferrase: Tyrosine-protein kinase; ATP-binding; Phosphorylation;
CC Polymorphism.
CC SIGNAL: 1 21 POTENTIAL.
CC CHAIN: 22 1255 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.
CC DOMAIN: 22 652 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM: 653 675 POTENTIAL.
CC DOMAIN: 676 1255 CYTOPLASMIC (POTENTIAL).
CC DOMAIN: 720 987 PROTEIN KINASE.
CC NP_BIND: 726 734 ATP (BY SIMILARITY).
CC BINDING: 753 753 ATP (BY SIMILARITY).
CC ACT_SITE: 845 845 BY SIMILARITY.
CC DISULFID: 195 204 BY SIMILARITY.
CC DISULFID: 199 212 BY SIMILARITY.
CC DISULFID: 220 227 BY SIMILARITY.
CC DISULFID: 224 235 BY SIMILARITY.
CC DISULFID: 236 244 BY SIMILARITY.
CC DISULFID: 240 252 BY SIMILARITY.
CC DISULFID: 255 264 BY SIMILARITY.
CC DISULFID: 268 295 BY SIMILARITY.
CC DISULFID: 289 311 BY SIMILARITY.
CC DISULFID: 315 331 BY SIMILARITY.
CC DISULFID: 334 338 BY SIMILARITY.
CC DISULFID: 511 520 BY SIMILARITY.
CC DISULFID: 515 528 BY SIMILARITY.
CC DISULFID: 531 540 BY SIMILARITY.
CC DISULFID: 544 560 BY SIMILARITY.

```

```

FT DISULFID 563 576 BY SIMILARITY.
FT DISULFID 567 584 BY SIMILARITY.
FT DISULFID 587 596 BY SIMILARITY.
FT DISULFID 600 623 BY SIMILARITY.
FT DISULFID 626 634 BY SIMILARITY.
FT DISULFID 630 642 BY SIMILARITY.
FT MOD_RES 1139 1139 BY SIMILARITY.
FT MOD_RES 1248 1248 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CAROHD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 187 187 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 571 571 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 654 654 I -> V.
FT VARIANT 654 654 /FTID=VAR_004077.
FT VARIANT 655 655 I -> V.
FT VARIANT 655 655 /FTID=VAR_004078.
FT CONFLICT 1170 1170 P -> A (IN REF. 2).
FT SEQUENCE 1255 AA; 137909 MW; 39E9DFDA04DCF962 CRC64;

Query Match 81.1%; Score 340; DB 1; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELALCRWGLLALLPPAASVCTGDMKRLPASPETHLMDLRLHYGCQVQGNL 60
DB 1 MELALCRWGLLALLPPAASVCTGDMKRLPASPETHLMDLRLHYGCQVQGNL 60
QY 61 ELTYLEPTNASLSFLDIOEVQGVYLAHNOVRQVPLQRLIRIVRGSQLFEDNYALAVLNDG 120
DB 61 ELTYLEPTNASLSFLDIOEVQGVYLAHNOVRQVPLQRLIRIVRGSQLFEDNYALAVLNDG 120
QY 121 DPLNNTPTVTGASPGGLRDLRLSELTEILKGVLLQNRNQLCYOPTILMKDIFHNNOQA 180
DB 121 DPLNNTPTVTGASPGGLRDLRLSELTEILKGVLLQNRNQLCYOPTILMKDIFHNNOQA 180
QY 121 DPLNNTPTVTGASPGGLRDLRLSELTEILKGVLLQNRNQLCYOPTILMKDIFHNNOQA 180
DB 121 DPLNNTPTVTGASPGGLRDLRLSELTEILKGVLLQNRNQLCYOPTILMKDIFHNNOQA 180
QY 181 LTLIDTNRSRACHPCSPMKSGSRGCESESDQSLTRYCAGGACRCKPLPTDCCHEQC 240
DB 181 LTLIDTNRSRACHPCSPMKSGSRGCESESDQSLTRYCAGGACRCKPLPTDCCHEQC 240
QY 241 AAGCTGPKHSDCLACIENHSGICELHCPALVTYNTDFESMPNDEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACIENHSGICELHCPALVTYNTDFESMPNDEGRYTFGASCVTACP 300
QY 301 YNVLSTDVSGCTLVCPRLHNOEVTAEEDGTQRCERCKSPCAR 340
DB 301 YNVLSTDVSGCTLVCPRLHNOEVTAEEDGTQRCERCKSPCAR 340

RESULT 2
ERR2_MESAU STANDARD; PRT; 1254 AA.
ID ERR2_MESAU
AC Q60553;
DR 15-DEC-1998 (Rel. 37, Created)
DR 15-DEC-1998 (Rel. 37, Last sequence update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE Receptor protein-tyrosine kinase erbb-2 precursor (EC 2.7.1.112)
DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2).
GN ERBB2 OR NEU.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxId=10036;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Nerve;
RX MEDLINE=94193007; PubMed=7908275;
RA Nakamura T., Ushijima T., Ishizaka Y., Nagao M., Arai M.,
RA Yamazaki Y., Ishikawa T.;
RA "Cloning and activation of the Syrian hamster neu proto-oncogene.";

```

RL Gene 140:251-255(1994).

CC -i- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,
 CC ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A
 CC POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-
 CC ALPHA AND AMPHIREGULIN (BY SIMILARITY).

CC -i- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.

CC -i- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS
 CC (POTENTIAL).

CC -i- SUBCELLULAR LOCATION: Type I membrane protein.

CC -i- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
 CC RESIDUES.

CC -i- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC -----

DR EMBL: D16295; BAA03801.1; -
 DR HSP; P11362; 1FGK.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR000719; Euk_kinase.
 DR InterPro: IPR002174; Furlin-like.
 DR InterPro: IPR001245; Tyr_kinase.
 DR InterPro: IPR004019; YLP_motif.
 DR Pfam: PF00069; kinase; 1.
 DR Pfam: PF00757; Furlin-like; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR Pfam: PF02757; YLP; 2.
 DR ProDom: PD000001; Euk_kinase; 1.
 DR SMART: SM00261; EU_3.
 DR SMART: SM00219; tyrc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 DR Transmembrane: Glycoprotein; Multigene family; Receptor; Signal;
 KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 KW Proto-oncogene; Disease mutation.

FT SIGNAL 1 21
 FT CHAIN 22 1254
 FT DOMAIN 22 652
 FT TRANSMEM 676 1254
 FT DOMAIN 158 368
 FT DOMAIN 472 644
 FT DOMAIN 720 987
 FT NP_BIND 726 734
 FT BINDING 753 753
 FT ACT_SITE 845 845
 FT DISULFID 195 204
 FT DISULFID 199 212
 FT DISULFID 236 244
 FT DISULFID 240 252
 FT DISULFID 255 264
 FT DISULFID 268 295
 FT DISULFID 299 311
 FT DISULFID 315 331
 FT DISULFID 334 338
 FT DISULFID 511 520
 FT DISULFID 515 528
 FT DISULFID 531 540
 FT DISULFID 544 560
 FT DISULFID 563 576
 FT DISULFID 567 584
 FT DISULFID 587 596
 FT DISULFID 600 623
 FT DISULFID 626 634
 FT DISULFID 630 642
 FT MOD_RES 1139 1139

FT MOD_RES 1247 1247
 FT CARBOHYD 68 68
 FT CARBOHYD 125 125
 FT CARBOHYD 187 187
 FT CARBOHYD 259 259
 FT CARBOHYD 530 530
 FT CARBOHYD 571 571
 FT CARBOHYD 629 629
 FT VARIANT 658 658
 FT VARIANT 659 659
 SQ SEQUENCE 1254 AA; 138252 MW; 974C3791C21F2BE1 CRC64;

Query Match 16.0%; Score 67; DB 1; Length 1254;
 Best Local Similarity 100.0%; Pred. No. 1.6e-59;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 LPDCCHEQCAAGCTGPKHSDCIACLFHNSGICELHCPALVYNTDFESMNPREGRYT 290
 DB 231 LPDCCHEQCAAGCTGPKHSDCIACLFHNSGICELHCPALVYNTDFESMNPREGRYT 290
 QY 291 FGASCVT 297
 DB 291 FGASCVT 297

RESULT 3
 ERB2_RAT
 ID ERB2_RAT STANDARD; PRT; 1257 AA.
 AC P06494;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)
 DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Epidermal growth factor
 DE receptor-related protein).
 GN ERB2 OR NEU.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Neuroblastoma;
 RX MEDLINE=86118662; PubMed=3945311;
 RA Bargman C.I., Hung M.-C., Weinberg R.A.;
 RT "The neu oncogene encodes an epidermal growth factor receptor-related
 RT protein.";
 RL Nature 319:226-230(1986).
 RN [2]
 RP SEQUENCE OF 852-905 FROM N.A.
 RC TISSUE=sciatic nerve;
 RX MEDLINE=9122560; PubMed=2025425;
 RA Lai C., Lemke G.;
 RT "An extended family of protein-tyrosine kinase genes differentially
 RT expressed in the vertebrate nervous system.";
 RL Neuron 6:691-704(1991).
 RN [3]
 RP STRUCTURE BY NMR OF 650-668.
 RX MEDLINE=9215181; PubMed=1346763;
 RA Gullick W.J., Bottomley A.C., Lofes F.J., Doak D.G., Mulvey D.,
 RA Newman R., Crumpton M.J., Sternberg M.J.E., Campbell I.D.;
 RT "Three dimensional structure of the transmembrane region of the proto-
 RT oncogenic and oncogenic forms of the neu protein.";
 RL EMBO J. 11:43-48(1992).

CC -i- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,
 CC ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A
 CC POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-
 CC ALPHA AND AMPHIREGULIN.

CC -i- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.

CC -i- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS.
 CC THE CONSTITUTIVELY ACTIVATED ONCOGENIC VARIANT FORMS A HOMODIMER.

CC -i- SUBCELLULAR LOCATION: Type I membrane protein.

CC	-1	PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES (BY SIMILARITY).	
CC	-1	SIMILARITY: BELONGS TO THE EGFR RECEPTOR FAMILY.	
CC			
CC		This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@sib-sib.ch).	
CC			
DR	EMBL: X03362; CAA27059.1; ALT_INIT.		
DR	PIR: A24562; TVRTNU.		
DR	HSSP: P13362; IEKG.		
DR	InterPro: IPR000494; EGFR_L_domain.		
DR	InterPro: IPR000719; Euk_kinase.		
DR	InterPro: IPR002174; Furin-like.		
DR	InterPro: IPR001245; Tyr_kinase.		
DR	InterPro: IPR004019; YLP_motif.		
DR	Pfam: PF00069; Furinase; 1.		
DR	Pfam: PF00757; Furin-like; 1.		
DR	Pfam: PF01030; Recep_L_domain; 2.		
DR	Pfam: PF02757; YLP; 2.		
DR	ProDom: PD000001; Euk_kinase; 1.		
DR	SMART: SMO0261; Fu; 3.		
DR	SMART: SMO0219; TYRK; 1.		
DR	PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.		
DR	PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.		
DR	PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.		
KW	Transmembrane; Glycoprotein; Multigene family; Receptor; Signal; Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation; Proto-oncogene; Disease mutation.		
KW	Proto-oncogene; Disease mutation.		
FT	SIGNAL	1	POTENTIAL.
FT	CHAIN	22	RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.
FT	DOMAIN	22	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	655	POTENTIAL.
FT	DOMAIN	678	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	159	CYS-RICH.
FT	DOMAIN	473	CYS-RICH.
FT	DOMAIN	722	PROTEIN KINASE.
FT	NP_BIND	728	ATP (BY SIMILARITY).
FT	BINDING	755	ATP (BY SIMILARITY).
FT	ACT_SITE	847	BY SIMILARITY.
FT	DISULFID	196	BY SIMILARITY.
FT	DISULFID	200	BY SIMILARITY.
FT	DISULFID	221	BY SIMILARITY.
FT	DISULFID	225	BY SIMILARITY.
FT	DISULFID	237	BY SIMILARITY.
FT	DISULFID	241	BY SIMILARITY.
FT	DISULFID	256	BY SIMILARITY.
FT	DISULFID	269	BY SIMILARITY.
FT	DISULFID	300	BY SIMILARITY.
FT	DISULFID	316	BY SIMILARITY.
FT	DISULFID	335	BY SIMILARITY.
FT	DISULFID	513	BY SIMILARITY.
FT	DISULFID	517	BY SIMILARITY.
FT	DISULFID	533	BY SIMILARITY.
FT	DISULFID	546	BY SIMILARITY.
FT	DISULFID	565	BY SIMILARITY.
FT	DISULFID	569	BY SIMILARITY.
FT	DISULFID	589	BY SIMILARITY.
FT	DISULFID	602	BY SIMILARITY.
FT	DISULFID	628	BY SIMILARITY.
FT	DISULFID	632	BY SIMILARITY.
FT	MOD_RES	1141	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	MOD_RES	1250	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	CARBOHYD	68	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	188	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	260	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	532	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	573	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	631	N-LINKED (GLCNAC. . .) (POTENTIAL).

FT	SEQUENCE	1257 AA;	661	661	V -E (IN ONCOGENIC NEU).	FT	SEQUENCE	1257 AA;	661	661	V -E (IN ONCOGENIC NEU).
SO	SEQUENCE	1257 AA;	138831 MM;	6129264583011402	CRC64;	SO	SEQUENCE	1257 AA;	138831 MM;	6129264583011402	CRC64;
Query Match		12.4%;	Score 52;	DB 1;	Length 1257;	Query Match		12.4%;	Score 52;	DB 1;	Length 1257;
Best Local Similarity		100.0%;	Pred. No. 2,7e-44;			Best Local Similarity		100.0%;	Pred. No. 2,7e-44;		
Matches	52;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;		
Oy	231	LPTDCHQCAAGCTGPKHSDCLALFHNSHSGICELHCALVTYNTDPRESM	282			Oy	231	LPTDCHQCAAGCTGPKHSDCLALFHNSHSGICELHCALVTYNTDPRESM	282		
Db	232	LPTDCHQCAAGCTGPKHSDCLALFHNSHSGICELHCALVTYNTDPRESM	283			Db	232	LPTDCHQCAAGCTGPKHSDCLALFHNSHSGICELHCALVTYNTDPRESM	283		
RESULT 4						RESULT 4					
RNF_E_VIBCH		STANDARD.	PRT;	230 AA.		RNF_E_VIBCH		STANDARD.	PRT;	230 AA.	
AC	Q9KT91;					AC	Q9KT91;				
DT	15-JUN-2002 (Rel. 41, Created)					DT	15-JUN-2002 (Rel. 41, Created)				
DT	15-JUN-2002 (Rel. 41, Last sequence update)					DT	15-JUN-2002 (Rel. 41, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)					DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Electron transport complex protein infE.					DE	Electron transport complex protein infE.				
GN	RNF_E_VIBCH					GN	RNF_E_VIBCH				
OS	Vibrio cholerae.					OS	Vibrio cholerae.				
OC	Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.					OC	Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.				
OX	NCBI_TaxID=666;					OX	NCBI_TaxID=666;				
RN	[1]					RN	[1]				
RP	SEQUENCE FROM N.A.					RP	SEQUENCE FROM N.A.				
RC	STRAIN=El TOR N16961 / Serotype O1;					RC	STRAIN=El TOR N16961 / Serotype O1;				
RX	MEDLINE=20406833; PubMed=10952301;					RX	MEDLINE=20406833; PubMed=10952301;				
RA	Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,					RA	Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,				
RA	Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,					RA	Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,				
RA	Gill S.R., Nelson K.E., Read T.D., Jettelin H., Richardson D.,					RA	Gill S.R., Nelson K.E., Read T.D., Jettelin H., Richardson D.,				
RA	Ernolova M.D., Vanathavan J., Bass S., Qin H., Dragol I., Sellers P.,					RA	Ernolova M.D., Vanathavan J., Bass S., Qin H., Dragol I., Sellers P.,				
RA	McDonald L., Uteerbeck T., Fleischmann R.D., Nierman W.C., White O.,					RA	McDonald L., Uteerbeck T., Fleischmann R.D., Nierman W.C., White O.,				
RA	Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,					RA	Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,				
RA	Fraser C.M.;					RA	Fraser C.M.;				
RT	"DNA sequence of both chromosomes of the cholera pathogen Vibrio					RT	"DNA sequence of both chromosomes of the cholera pathogen Vibrio				
RL	cholerae."					RL	cholerae."				
CC	Nature 406:477-483(2000).					CC	Nature 406:477-483(2000).				
CC	-I- FUNCTION: May be part of a membrane complex involved in electron					CC	-I- FUNCTION: May be part of a membrane complex involved in electron				
CC	transport (By similarity).					CC	transport (By similarity).				
CC	-I- SUBUNIT: Composed of at least six subunits; rnfA, rnfB, rnfC,					CC	-I- SUBUNIT: Composed of at least six subunits; rnfA, rnfB, rnfC,				
CC	rnfD, rnfE and rnfG (By similarity)					CC	rnfD, rnfE and rnfG (By similarity)				
CC	-I- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane					CC	-I- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane				
CC	(Potential).					CC	(Potential).				
CC	-I- SIMILARITY: BELONGS TO THE NORDE/RNFAE FAMILY.					CC	-I- SIMILARITY: BELONGS TO THE NORDE/RNFAE FAMILY.				
CC	-----					CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation					CC	between the Swiss Institute of Bioinformatics and the EMBL outstation				
CC	at the European Bioinformatics Institute. There are no restrictions on its					CC	at the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way					CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial					CC	modified and this statement is not removed. Usage by and for commercial				
CC											

```

RESULT 5
EGFR_CHICK
ID EGFR_CHICK STANDARD: PRT; 703 AA.
AC P1387;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (CER)
DE (Fragment).
GN EGFR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=88261272; PubMed=3260329;
RA Lax I., Johnson A., Howk R., Sap J., Bellot F., Winkler M.,
RA Ullrich A., Vennstrom B., Schlessinger J., Glavo D.;
RT "Chicken epidermal growth factor (EGF) receptor: cDNA cloning,
RT expression in mouse cells, and differential binding of EGF and
RT transforming growth factor alpha."
RL Mol. Cell. Biol. 8:1970-1978(1988).
CC -1- FUNCTION: THE EGF RECEPTOR MEDIATES THE BIOLOGICAL SIGNAL OF EGF.
CC AND ALSO OF TGF-ALPHA, AMPHIREGULIN, HEPARIN-BINDING EGF, GP30 AND
CC VACCINIA VIRUS GROWTH FACTOR (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- MISCELLANEOUS: Binding of EGF to the receptor leads to
CC dimerization, internalization of the EGF-receptor complex,
CC induction of the tyrosine kinase activity, stimulation of cell DNA
CC synthesis, and cell proliferation.
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
DR EMBL: M20386; AAA8760.1;
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR001245; Tyr-kinase.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR SMART: SM00261; FU; 4.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; PARTIAL.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; PARTIAL.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; PARTIAL.
KM Transmembrane: Glycoprotein; Receptor; Signal; Transferase;
KM Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT SIGNAL 1 30
FT CHAIN 31 >703
FT DOMAIN 31 654 EPIDERMAL GROWTH FACTOR RECEPTOR.
FT TRANSMEM 655 667 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 668 >703 POTENTIAL.
FT DISULFID 197 206 CYTOPLASMIC (POTENTIAL).
FT DISULFID 201 206 BY SIMILARITY.
FT DISULFID 222 230 BY SIMILARITY.
FT DISULFID 226 230 BY SIMILARITY.
FT DISULFID 239 247 BY SIMILARITY.
FT DISULFID 243 255 BY SIMILARITY.
FT DISULFID 258 267 BY SIMILARITY.
FT DISULFID 271 298 BY SIMILARITY.
FT DISULFID 302 314 BY SIMILARITY.

```

```

FT DISULFID 318 333 BY SIMILARITY.
FT DISULFID 336 340 BY SIMILARITY.
FT DISULFID 513 522 BY SIMILARITY.
FT DISULFID 517 530 BY SIMILARITY.
FT DISULFID 533 542 BY SIMILARITY.
FT DISULFID 546 562 BY SIMILARITY.
FT DISULFID 565 581 BY SIMILARITY.
FT DISULFID 569 589 BY SIMILARITY.
FT DISULFID 592 601 BY SIMILARITY.
FT DISULFID 605 627 BY SIMILARITY.
FT DISULFID 630 638 BY SIMILARITY.
FT DISULFID 634 646 BY SIMILARITY.
FT CARBOHYD 134 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 359 359 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 420 420 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 573 573 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 578 578 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 613 613 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 633 633 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 648 648 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 703
SQ SEQUENCE 703 AA; 77427 MW; AFD2DE1B735A690 CRC64;

Query Match          2.1%; Score 9; DB 1; Length 703;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 239 QCAAGCTGP 247
   |||||
Db 242 QCAAGCTGP 250

RESULT 6
EGFR_HUMAN
ID EGFR_HUMAN STANDARD: PRT; 1210 AA.
AC P00533; P06268; 014225; Q9UMD7; Q9UMB8; Q9UMK5; Q92795; 000732;
AC 000688; Q9R2S2; Q9H2C9; Q9GZX1; Q9H3C9;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (Receptor
DE protein-tyrosine kinase ErbB-1).
GN EGFR OR ERBB1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=84219729; PubMed=6328312;
RA Ullrich A., Coussens L., Hayflick J.S., Dull T.J., Gray A., Tam A.W.,
RA Lee J., Yarden Y., Libermann T.A., Schlessinger J., Downward J.,
RA Mayes E.L.V., Whittle N., Waterfield M.D., Seeburg P.H.;
RT "Human epidermal growth factor receptor cDNA sequence and aberrant
RT expression of the amplified gene in A431 epidermoid carcinoma cells."
RT Nature 309:418-425(1984).
RN (2)
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX TISSUE-Placenta;
RX MEDLINE=95382957; PubMed=7654368;
RA Ilekis J.V., Stark B.C., Scoccia B.;
RT "Possible role of variant RNA transcripts in the regulation of
RT epidermal growth factor receptor expression in human placenta."
RL Mol. Reprod. Dev. 41:149-156(1995).
RN (3)
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX TISSUE-Placenta;
RX MEDLINE=97078686; PubMed=8918811;
RA Reiter J.L., Mahle N.J.;
RT "A 1.8 kb alternative transcript from the human epidermal growth

```

RT factor receptor gene encodes a truncated form of the receptor.";
 RL Nucleic Acids Res. 24:4050-4056(1996).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE-Placenta;
 RX MEDLINE=97256547; PubMed=9103388;
 RT Ilekis J.V., Garliti J., Niederberger C., Scoccia B.;
 RA "Expression of a truncated epidermal growth factor receptor-like
 RT protein (TEGFR) in ovarian cancer.";
 RL Gynecol. Oncol. 65:36-41(1997).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
 RC TISSUE-Placenta;
 RX MEDLINE=21100872; PubMed=11161793;
 RT Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,
 RA Schehl Sinclair C., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,
 RA Balasubramanian S., Crossley T.D., Magnuson T.R., James C.D.,
 RA Mailhe N.J.;
 RT "Comparative genomic sequence analysis and isolation of human and
 RT mouse alternative EGF transcripts encoding truncated receptor
 RT isoforms.";
 RL Genomics 71:1-20(2001).
 RN [6]
 RP SEQUENCE OF 575-687 FROM N.A.
 RA Reiter J.L., Threadgill D.W., Danielson A.J., Schehl C.M.,
 RA Lampland A.L., Balasubramanian S., Crossley T.O., Magnuson T.R.,
 RA Mailhe N.J.;
 RT "Human and mouse alternative EGF transcripts encoding only the
 RT extracellular domain of the receptor";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 713-924 FROM N.A.
 RX MEDLINE=84196372; PubMed=6326261;
 RA Lin C.R., Chen W.S., Kruijer W., Stolarsky L.S., Weber W.,
 RA Evans R.M., Verma I.M., Gill G.N., Rosenfeld M.G.;
 RT "Expression cloning of human EGF receptor complementary DNA: gene
 RT amplification and three related messenger RNA products in A431
 RT cells.";
 RL Science 224:843-848(1984).
 RN [8]
 RP SEQUENCE OF 150-962 FROM N.A.
 RX MEDLINE=84245835; PubMed=6330563;
 RA Xu Y.H., Ishii S., Clark A.J.L., Sullivan M., Wilson R.K., Ma D.P.,
 RA Roe B.A., Merlino G.T., Pastan I.;
 RT "Human epidermal growth factor receptor cDNA is homologous to a
 RT variety of RNAs overproduced in A431 carcinoma cells.";
 RL Nature 309:806-810(1984).
 RN [9]
 RP SEQUENCE OF 1028-1210 FROM N.A.
 RX MEDLINE=85046483; PubMed=6093780;
 RA Simmen F.A., Gope M.L., Schultz T.Z., Wright D.A., Carpenter G.,
 RA O'Malley B.W.;
 RT "Isolation of an evolutionarily conserved epidermal growth factor
 RT receptor cDNA from human A431 carcinoma cells.";
 RL Biochem. Biophys. Res. Commun. 124:123-132(1984).
 RN [10]
 RP SEQUENCE OF 1-29 FROM N.A.
 RX MEDLINE=88217333; PubMed=3329716;
 RA Haley J.D., Whittle N., Bennett P., Kinchington D., Ullrich A.,
 RA Waterfield M.D.;
 RT "The human EGF receptor gene: structure of the 110 kb locus and
 RT identification of sequences regulating its transcription.";
 RL Oncogene Res. 1:375-396(1987).
 RN [11]
 RP SEQUENCE OF 1-29 FROM N.A.
 RX MEDLINE=91107677; PubMed=1988448;
 RA Haley J.D., Waterfield M.D.;
 RT "Contributory effects of de novo transcription and premature
 RT transcript termination in the regulation of human epidermal growth
 RT factor receptor proto-oncogene RNA synthesis.";
 RL J. Biol. Chem. 266:1746-1753(1991).
 RN [12]
 RP SEQUENCE OF 1-29 FROM N.A.
 RX MEDLINE=85270438; PubMed=2991899;
 RA Ishii S., Xu Y.H., Stratton R.H., Roe B.A., Merlino G.T., Pastan I.;
 RT "Characterization and sequence of the promoter region of the human
 RT epidermal growth factor receptor gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:4920-4924(1985).
 RN [13]
 RP SEQUENCE OF 540.
 RA Kohda D.;
 RT Submitted (SEP-1997) to the SWISS-PROT data bank.
 RN [14]
 RP RECEPTOR ACTIVITY.
 RX MEDLINE=84191554; PubMed=6325948;
 RA Mroczkowski B., Mosig G., Cohen S.;
 RT "ATP-stimulated interaction between epidermal growth factor receptor
 RT and supercoiled DNA.";
 RL Nature 309:270-273(1984).
 RN [15]
 RP PHOSPHORYLATION.
 RX MEDLINE=89278137; PubMed=2543678;
 RA Margolis B.L., Lax I., Kris R., Dombalagian M., Honegger A.M.,
 RA Howk R., Givol D., Ullrich A., Schlessinger J.;
 RT "All autophosphorylation sites of epidermal growth factor (EGF)
 RT receptor and HER2/neu are located in their carboxyl-terminal tails.
 RT Identification of a novel site in EGF receptor.";
 RL J. Biol. Chem. 264:10667-10671(1989).
 RN [16]
 RP CARBOHYDRATE-LINKAGE SITES ASN-128; ASN-175; ASN-413; ASN-444 AND
 RP ASN-528.
 RX MEDLINE=96398132; PubMed=8962717;
 RA Smith K.D., Davies M.J., Bailey D., Renouf D.V., Hounsell E.F.;
 RT "Analysis of the glycosylation patterns of the extracellular domain of
 RT the epidermal growth factor receptor expressed in Chinese hamster
 RT ovary fibroblasts.";
 RL Growth Factors 13:121-132(1996).
 RN [17]
 RP CARBOHYDRATE-LINKAGE SITES ASN-56; ASN-352; ASN-361; ASN-568 AND
 RP ASN-603.
 RX MEDLINE=20198209; PubMed=10731668;
 RA Sato C., Kim J.-H., Abe Y., Saito K., Yokoyama S., Kohda D.;
 RT "Characterization of the N-oligosaccharides attached to the atypical
 RT Asn-X-Cys sequence of recombinant human epidermal growth factor
 RT receptor.";
 RL J. Biochem. 127:65-72(2000).
 RN [18]
 RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
 RX MEDLINE=98225196; PubMed=9556602;
 RA Abe Y., Otake M., Inagaki F., Lax I., Schlessinger J., Kohda D.;
 RT "Disulfide bond structure of human epidermal growth factor receptor.";
 RL J. Biol. Chem. 273:11150-11157(1998).
 RN [19]
 RP REVIEW.
 RX MEDLINE=87297456; PubMed=3039909;
 RA Carpenter G.;
 RT "Receptors for epidermal growth factor and other polypeptide
 RT mitogens.";
 RL Annu. Rev. Biochem. 56:881-914(1987).
 CC -1- FUNCTION: Receptor for EGF, but also for other members of the EGF
 CC family, as TGF-alpha, amphiregulin, betacellulin, heparin-binding
 CC EGF-like growth factor, GP30 and vaccinia virus growth factor. It
 CC is involved in the control of cell growth and differentiation.
 CC -1- FUNCTION: Isoform 2/truncated isoform may act as an antagonist.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Isoform 2 is
 CC secreted.
 CC -1- ALTERNATIVE PRODUCTS: 4 isoforms; 1/p170 (shown here), 2/p60/
 CC truncated isoform/TEGFR, 3/p110 and 4; are produced by
 CC alternative splicing.
 CC -1- TISSUE SPECIFICITY: Expressed in placenta. Isoform 2 is also
 CC expressed in ovarian cancers.
 CC -1- MISCELLANEOUS: Binding of EGF to the receptor leads to
 CC dimerization, internalization of the EGF-receptor complex,
 CC induction of the tyrosine kinase activity, stimulation of cell DNA


```

FT DISULFID 295 307 BY SIMILARITY.
FT DISULFID 311 326 BY SIMILARITY.
FT DISULFID 329 333 BY SIMILARITY.
FT DISULFID 506 515 BY SIMILARITY.
FT DISULFID 510 523 BY SIMILARITY.
FT DISULFID 526 535 BY SIMILARITY.
FT DISULFID 539 555 BY SIMILARITY.
FT DISULFID 558 571 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
FT DISULFID 595 617 BY SIMILARITY.
FT DISULFID 620 628 BY SIMILARITY.
FT DISULFID 624 636 BY SIMILARITY.
FT MOD.RES 680 680 BY SIMILARITY.
FT MOD.RES 1092 1092 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
FT MOD.RES 1110 1110 PHOSPHORYLATION (BY SIMILARITY).
FT MOD.RES 1172 1172 PHOSPHORYLATION (BY SIMILARITY).
FT MOD.RES 1197 1197 PHOSPHORYLATION (BY SIMILARITY).
FT CARBOHYD 128 128 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 413 413 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 528 528 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 568 568 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 19 19 C -> S (IN REF. 2).
FT CONFLICT 539 539 C -> W (IN REF. 5).
FT CONFLICT 991 991 L -> F (IN REF. 4).
FT CONFLICT 1116 1117 HP -> DR (IN REF. 6).
SQ SEQUENCE 1210 AA; 134853 MW; 690E20D46DFD2F5 CRC64;

Query Match
Best Local Similarity 100.0%; Score 9; DB 1; Length 1210;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 239 QCAAGCTGP 247
Db 235 QCAAGCTGP 243

RESULT 8
DHCA_RAT STANDARD: PRT; 276 AA.
AC P47727;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Carboxyl reductase [NADPH] 1 (EC 1.1.1.184) (NADPH-dependent carboxyl
DE reductase 1).
GN CBRL OR CBR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Testis;
RA MEDLINE=95220378; PubMed=7705364;
RA Wernuth B., Meeder-Heinemann G., Ernst E.;
RT "Cloning and expression of carboxyl reductase from rat testis.";
RT Eur. J. Biochem. 228:473-479(1995).
CC -1- CATALYTIC ACTIVITY: R-CHOH-R' + NADP(+) = R-CO-R' + NADPH.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDS) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X84349; CAA59088.1; -.
CC EMBL; X95986; CAA65230.1; -.
CC HSSP; P50162; IAEI.
CC InterPro: IPR002198; ADH_short.
CC Pfam: PF00106; adh_short; 1.
CC PROSITE; P500061; ADH_SHORT; 1.
CC Oxidoreductase; NADP; Acetylation.
FT INIT_MET 0 0 BY SIMILARITY.
FT MOD.RES 1 1 ACETYLATION (BY SIMILARITY).
FT NP_BIND 9 33 NADP (BY SIMILARITY).
FT ACT_SITE 193 193 BY SIMILARITY.
SQ SEQUENCE 276 AA; 30447 MW; 99B9E77C5E2922AB CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 1; Length 276;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 LALLPPGA 20
Db 253 LALLPPGA 260

RESULT 9
YV30_MYCTU STANDARD: PRT; 463 AA.
AC 007035;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein RV3130C.
GN RV3130C OR MT3216 OR MTCY03A2.28 OR MTCY164.41c.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteriales; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tebala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skellon S., Squares S., Squares R.,
RA Sulston J.F., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RT Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bisphai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RT Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE UPF0089 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```


CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: 295150; CAB08399.1; -
DR EMBL: AE007137; AAK47554.1; -
DR TIGR: MT3216; -
DR TubercuList; RV3130c; -
DR InterPro: IPR004255; UPF0089.
DR Pfam: PF03007; UPF0089; 1.
KM Hypothetical protein: Complete proteome.
FT CONFLICT 268 268 Complete proteome.
SQ SEQUENCE 463 AA; 50720 MW; 48CC60A8670F1D6B CRC64;

Query Match 1.9%; Score 8; DB 1; Length 463;
Best local Similarity 100.0%; Pred. No. 5.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 PLOURLRV 102
DB 315 PLOURLRV 322

RESULT 10
NCAM2_MOUSE
ID NCAM2_MOUSE STANDARD; PRT; 725 AA.
AC P13594; Q61950;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Neural cell adhesion molecule 1, 120 kDa isoform precursor (N-CAM 120)
DE (NCAM-120).
GN NCAM1 OR NCAM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=87246524; PubMed=3595563;
RA Barthele D., Santoni M.-J., Wille M., Ruppert C., Caix J.-C.,
RT Hirsch M.-R., Fontecilla-Camps J.-C., Goridis C.;
RT "Isolation and nucleotide sequence of mouse NCAM cDNA that codes for
a Mr 79,000 polypeptide without a membrane-spanning region.";
RL EMBL J. 6:907-914(1987).
RN [2]
RP SEQUENCE OF 20-700 FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=89251563; PubMed=2721486;
RA Santoni M.-J., Barthele D., Vopper G., Boned A., Goridis C., Wille M.;
RT "Differential exon usage involving an unusual splicing mechanism
generates at least eight types of NCAM cDNA in mouse brain.";
RL EMBL J. 8:385-392(1989).
RN [3]
RP SEQUENCE OF 642-725 FROM N.A.
RC MEDLINE=88283628; PubMed=3396534;
RA Barbas J.A., Chaix J.C., Steinmetz M., Goridis C.;
RT "Differential splicing and alternative polyadenylation generates
distinct NCAM transcripts and proteins in the mouse.";
RL EMBL J. 7:625-632(1988).
RN [4]
RP SEQUENCE OF 20-36.
RX MEDLINE=86140120; PubMed=3512556;
RA Rougon G., Marshak D.R.;
RT "Structural and immunological characterization of the amino-terminal
domain of mammalian neural cell adhesion molecules.";
RL J. Biol. Chem. 261:3396-3401(1986).
CC -1- FUNCTION: THIS PROTEIN IS A CELL ADHESION MOLECULE INVOLVED IN
CC NEURON-NEURON ADHESION, NEURITE FASCICULATION, OUTGROWTH OF
CC NEURITES, ETC.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- ALTERNATIVE PRODUCTS: 3 isoforms; N-CAM 180 (AC P13595), N-CAM 140

CC (AC P13595) and N-CAM 120 (shown here); are produced by
CC alternative splicing.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -1- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: Y00051; CAA68263.1; -
DR EMBL: X15049; CAA33148.1; ALT_SEQ.
DR EMBL: X07195; CAA30173.1; -
DR PIR: A29673; IJMSNG.
DR MGD: MGI:97281; Ncam1.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003598; Ig_C2.
DR Pfam: PF00041; fn3; 2.
DR Pfam: PF00047; Ig; 5.
DR SMART: SM00060; FN3; 2.
DR SMART: SM00408; Igc2; 5.
KM Cell adhesion: Glycoprotein: Repeat: Alternative splicing:
KM Immunoglobulin domain: Signal; Heparin-binding; GPI-anchor.
FT CHAIN 1 19
FT 20 725
FT 103
FT 132 103
FT 132 196
FT 228 196
FT 323 196
FT 420 196
FT 487 196
FT 519 196
FT 625 196
FT 692 196
FT 152 196
FT 161 196
FT 165 196
FT 165 196
FT 189 196
FT 235 196
FT 330 196
FT 427 196
FT 480 196
FT 222 196
FT 316 196
FT 348 196
FT 424 196
FT 450 196
FT 479 196
FT 261 196
FT 268 196
FT 273 196
FT 354 196
FT 549 196
FT 572 196
FT 575 196
FT 589 196
FT 602 196
FT 657 196
SQ SEQUENCE 725 AA; 80296 MW; C2AE8BB461C8B2F CRC64;

Query Match 1.9%; Score 8; DB 1; Length 725;
Best local Similarity 100.0%; Pred. No. 7.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 322 VTRADSTQ 329
DB 98 VTRADSTQ 105

RESULT 11
NCAL_RAT


```

RESULT 14
ID Y164.ARCFU STANDARD: PRT: 104 AA.
AC 028415:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AF1864.
AF1864.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RA MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kierlavage A.R., Graham D.E., Kyriides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirschner E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -1- SIMILARITY: Strong, to A.aeolicus AQ_377.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AE000974; AAB89396.1; -.
DR TIGR: AF1864; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 104 AA; 11845 MW; DC80B36CF3F5C153 CRC64;

Query Match
Best Local Similarity 1.7%; Score 7; DB 1; Length 104;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 147 EILKGV 153
DB 95 EILKGV 101

RESULT 15
MADE_HUMAN
ID MADE_HUMAN STANDARD: PRT: 111 AA.
AC 000994:
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE p75NTR-associated cell death executor (Nerve growth factor receptor
DE associated protein 1) (Ovarian granulosa cell 13.0 kDa protein HGR74).
GN NGFRAP1 OR MADE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=91025550; PubMed=2171551;

```

```

RA Rapp G., Freudenstein J., Klaudiny J., Mucha J., Wempe F., Zimmer M.,
RA Scheit K.H.;
RT "Characterization of three abundant mRNAs from human ovarian
RT granulosa cells."
RT DNA Cell Biol. 9:479-485(1990).
RN [2]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=20298829; PubMed=10764727;
RA Mukai J., Hachiya T., Shoji-Hoshino S., Kimura M.T., Nadano D.,
RA Suvanto P., Hanaka T., Li Y., Irie S., Greene L.A., Sato T.A.;
RT "MADE, a p75NTR-associated cell death executor, is involved in signal
RT transduction mediated by the common neurotrophin receptor p75NTR."
RT J. Biol. Chem. 275:17566-17570(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strusberg R.;
RT Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: May be a signaling adaptor molecule involved in p75NTR-
CC mediated apoptosis induced by NGF. May play an important role in
CC the pathogenesis of neurogenetic diseases.
CC -1- SUBUNIT: Binds to the DEATH domain of p75NTR/NGFR.
CC -1- TISSUE SPECIFICITY: FOUND IN OVARIAN GRANULOSA CELLS, TESTIS,
CC PROSTATE AND SEMINAL VESICLE TISSUE.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M38188; AAA63232.1; -.
DR EMBL: AF187064; AAF75129.1; -.
DR EMBL: BC003190; AAH03190.1; -.
DR PIR: C35826; C35826.
DR Genew: HGNC:13388; NGFRAP1.
DR MIM: 300361; -.
DR InterPro: IPR001230; Prenyl-site.
KW Apoptosis.
SQ SEQUENCE 111 AA; 12958 MW; 29AA0573282C933E CRC64;

Query Match
Best Local Similarity 1.7%; Score 7; DB 1; Length 111;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 LRELQLR 143
DB 81 LRELQLR 87

Search completed: March 4, 2003, 12:45:38
Job time : 30.6064 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:42:02 ; Search time 46.2751 Seconds
(without alignments)
1865.663 Million cell updates/sec

Title: US-09-234-208b-2

Perfect score: 419
Sequence: 1 MELALCRWGLLALLPPGA.....VGRGPDPAHVAVNLSEYEG 419

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPREMBL_21.*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp-invertebrate:*
6: sp-mammal:*
7: sp-mhc:*
8: sp-organelle:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp-unclassified:*
15: sp-virus:*
16: sp-bacteriap:*
17: sp-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	345	82.3	419	4	Q9UK79
2	67	16.0	149	6	Q9BG66
3	66	15.8	1259	6	Q9BG66
4	9	2.1	478	11	Q18735
5	9	2.1	527	13	Q9ES60
6	9	2.1	599	13	Q9PSH2
7	9	2.1	643	11	Q9ERV6
8	9	2.1	655	11	Q9WVFS
9	9	2.1	1209	11	Q9QX70
10	9	2.1	1210	11	Q9EP98
11	8	1.9	48	11	Q63365
12	8	1.9	63	11	Q9QZ16
13	8	1.9	63	11	Q9QZ15
14	8	1.9	217	5	Q9N6W6
15	8	1.9	277	11	Q08558
16	8	1.9	277	11	Q924V3

17	8	1.9	280	13	Q8UVJ2
18	8	1.9	341	16	Q91072
19	8	1.9	348	16	Q97S36
20	8	1.9	481	10	Q945M4
21	8	1.9	520	16	Q66518
22	8	1.9	583	16	Q99RE2
23	8	1.9	605	11	Q921P2
24	8	1.9	865	10	Q22774
25	8	1.9	879	10	Q41010
26	8	1.9	1436	17	Q96ZJ5
27	8	1.9	1469	10	Q9LKR1
28	8	1.9	1503	10	Q81283
29	7	1.7	9	2	Q93LE4
30	7	1.7	27	4	Q9HD19
31	7	1.7	39	5	Q8WCU2
32	7	1.7	61	10	Q8S9S5
33	7	1.7	67	12	Q84264
34	7	1.7	114	11	Q9D1N5
35	7	1.7	117	2	Q8VMI3
36	7	1.7	120	11	Q9JIT2
37	7	1.7	121	16	Q8YZW8
38	7	1.7	124	11	Q9CWN9
39	7	1.7	124	11	Q9WT29
40	7	1.7	146	15	Q8UJ38
41	7	1.7	146	15	Q8UJ36
42	7	1.7	147	11	Q9D0S2
43	7	1.7	150	6	Q9BG64
44	7	1.7	162	10	Q9EPD4
45	7	1.7	163	11	Q9JLY9

ALIGNMENTS

RESULT 1
Q9UK79 PRELIMINARY; PRT; 419 AA.
AC Q9UK79;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Herstatin.
GN HER-2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99415951; PubMed=10485918;
RA Doherty J.K., Bond C., Jardim A., Adelman J.P., Clinton G.M.;
RT "The HER-2/neu Receptor tyrosine kinase gene encodes a secreted
RT autoinhibitor.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:10869-10874(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Doherty J.K., Clinton G.M., Adelman J.P., Evans A.J., Henner W.D.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF177761; AAD56009.2; -
DR InterPro: IPR00494; EGFRL_domain.
DR InterPro: IPR002174; Furin-like.
DR pfam: PF00757; Furin-like; 1.
DR pfam: PF01030; Recep_L_domain; 1.
DR SMART: SM00261; FU; 1.
SQ SEQUENCE 419 AA; 45472 MW; FEC1BE347E2D030C CRC64;

Query Match 82.3%; Score 345; DB 4; Length 419;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MELALCRWGLLALLPPGAASVQVCTGDMKRLRPASPEFHLDMLRHLYGCGQVQGNL 60
Db 1 MELALCRWGLLALLPPGAASVQVCTGDMKRLRPASPEFHLDMLRHLYGCGQVQGNL 60

```

OY 61 ELTYLPTNLSLFLQDIOEVOGYVLLAHNOVRQVPLQRLRIYKGTOLFEEDNVALAVLNG 120
    |||||||
DB 61 ELTYLPTNLSLFLQDIOEVOGYVLLAHNOVRQVPLQRLRIYKGTOLFEEDNVALAVLNG 120
OY 121 DPLNNTPTVTGASPGGLRELQRLSLREILKGVLIQRPOLCYDTITLWKDFHKNNOIA 180
    |||||||
DB 121 DPLNNTPTVTGASPGGLRELQRLSLREILKGVLIQRPOLCYDTITLWKDFHKNNOIA 180
OY 181 LTLIDNRSRACHPCSPMCKSGRCWESSSDCQSLRTVCAGGACARCKGPLPDCCHQC 240
    |||||||
DB 181 LTLIDNRSRACHPCSPMCKSGRCWESSSDCQSLRTVCAGGACARCKGPLPDCCHQC 240
OY 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDFESMPNPGRYTFGASCYTAC 300
    |||||||
DB 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDFESMPNPGRYTFGASCYTAC 300
OY 301 YNLTSTDVSGCTLVCPDLHNOVTAEDGTQRCCKSPCARGTHSL 345
    |||||||
DB 301 YNLTSTDVSGCTLVCPDLHNOVTAEDGTQRCCKSPCARGTHSL 345

```

RESULT 2

```

OY 61 ELTYLPTNLSLFLQDIOEVOGYVLLAHNOVRQVPLQRLRIYKGTOLFEEDNVALAVLNG 120
    |||||||
DB 61 ELTYLPTNLSLFLQDIOEVOGYVLLAHNOVRQVPLQRLRIYKGTOLFEEDNVALAVLNG 120
OY 121 DPLNNTPTVTGASPGGLRELQRLSLREILKGVLIQRPOLCYDTITLWKDFHKNNOIA 180
    |||||||
DB 121 DPLNNTPTVTGASPGGLRELQRLSLREILKGVLIQRPOLCYDTITLWKDFHKNNOIA 180
OY 181 LTLIDNRSRACHPCSPMCKSGRCWESSSDCQSLRTVCAGGACARCKGPLPDCCHQC 240
    |||||||
DB 181 LTLIDNRSRACHPCSPMCKSGRCWESSSDCQSLRTVCAGGACARCKGPLPDCCHQC 240
OY 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDFESMPNPGRYTFGASCYTAC 300
    |||||||
DB 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDFESMPNPGRYTFGASCYTAC 300
OY 301 YNLTSTDVSGCTLVCPDLHNOVTAEDGTQRCCKSPCARGTHSL 345
    |||||||
DB 301 YNLTSTDVSGCTLVCPDLHNOVTAEDGTQRCCKSPCARGTHSL 345

```

Query Match 16.0%; Score 67; DB 6; Length 149;

Best local Similarity 100.0%; Pred. No. 2e-60; Mismatches 0; Indels 0; Gaps 0;

```

OY 231 LPTDCHEQCAAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDFESMPNPGRYT 290
    |||||||
DB 74 LPTDCHEQCAAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDFESMPNPGRYT 133
OY 291 FGASCVT 297
    |||||||
DB 134 FGASCVT 140

```

RESULT 3

```

OY 61 ELTYLPTNLSLFLQDIOEVOGYVLLAHNOVRQVPLQRLRIYKGTOLFEEDNVALAVLNG 120
    |||||||
DB 61 ELTYLPTNLSLFLQDIOEVOGYVLLAHNOVRQVPLQRLRIYKGTOLFEEDNVALAVLNG 120
OY 121 DPLNNTPTVTGASPGGLRELQRLSLREILKGVLIQRPOLCYDTITLWKDFHKNNOIA 180
    |||||||
DB 121 DPLNNTPTVTGASPGGLRELQRLSLREILKGVLIQRPOLCYDTITLWKDFHKNNOIA 180
OY 181 LTLIDNRSRACHPCSPMCKSGRCWESSSDCQSLRTVCAGGACARCKGPLPDCCHQC 240
    |||||||
DB 181 LTLIDNRSRACHPCSPMCKSGRCWESSSDCQSLRTVCAGGACARCKGPLPDCCHQC 240
OY 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDFESMPNPGRYTFGASCYTAC 300
    |||||||
DB 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDFESMPNPGRYTFGASCYTAC 300
OY 301 YNLTSTDVSGCTLVCPDLHNOVTAEDGTQRCCKSPCARGTHSL 345
    |||||||
DB 301 YNLTSTDVSGCTLVCPDLHNOVTAEDGTQRCCKSPCARGTHSL 345

```

```

OY 61 ELTYLPTNLSLFLQDIOEVOGYVLLAHNOVRQVPLQRLRIYKGTOLFEEDNVALAVLNG 120
    |||||||
DB 61 ELTYLPTNLSLFLQDIOEVOGYVLLAHNOVRQVPLQRLRIYKGTOLFEEDNVALAVLNG 120
OY 121 DPLNNTPTVTGASPGGLRELQRLSLREILKGVLIQRPOLCYDTITLWKDFHKNNOIA 180
    |||||||
DB 121 DPLNNTPTVTGASPGGLRELQRLSLREILKGVLIQRPOLCYDTITLWKDFHKNNOIA 180
OY 181 LTLIDNRSRACHPCSPMCKSGRCWESSSDCQSLRTVCAGGACARCKGPLPDCCHQC 240
    |||||||
DB 181 LTLIDNRSRACHPCSPMCKSGRCWESSSDCQSLRTVCAGGACARCKGPLPDCCHQC 240
OY 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDFESMPNPGRYTFGASCYTAC 300
    |||||||
DB 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDFESMPNPGRYTFGASCYTAC 300
OY 301 YNLTSTDVSGCTLVCPDLHNOVTAEDGTQRCCKSPCARGTHSL 345
    |||||||
DB 301 YNLTSTDVSGCTLVCPDLHNOVTAEDGTQRCCKSPCARGTHSL 345

```

Query Match

Best local Similarity 100.0%; Pred. No. 1.5e-58; Mismatches 0; Indels 0; Gaps 0;

```

OY 232 PDCCHQCAAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDFESMPNPGRYT 291
    |||||||
DB 232 PDCCHQCAAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDFESMPNPGRYT 291
OY 292 GASCVT 297
    |||||||
DB 292 GASCVT 297

```

RESULT 4

```

OY 61 ELTYLPTNLSLFLQDIOEVOGYVLLAHNOVRQVPLQRLRIYKGTOLFEEDNVALAVLNG 120
    |||||||
DB 61 ELTYLPTNLSLFLQDIOEVOGYVLLAHNOVRQVPLQRLRIYKGTOLFEEDNVALAVLNG 120
OY 121 DPLNNTPTVTGASPGGLRELQRLSLREILKGVLIQRPOLCYDTITLWKDFHKNNOIA 180
    |||||||
DB 121 DPLNNTPTVTGASPGGLRELQRLSLREILKGVLIQRPOLCYDTITLWKDFHKNNOIA 180
OY 181 LTLIDNRSRACHPCSPMCKSGRCWESSSDCQSLRTVCAGGACARCKGPLPDCCHQC 240
    |||||||
DB 181 LTLIDNRSRACHPCSPMCKSGRCWESSSDCQSLRTVCAGGACARCKGPLPDCCHQC 240
OY 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDFESMPNPGRYTFGASCYTAC 300
    |||||||
DB 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDFESMPNPGRYTFGASCYTAC 300
OY 301 YNLTSTDVSGCTLVCPDLHNOVTAEDGTQRCCKSPCARGTHSL 345
    |||||||
DB 301 YNLTSTDVSGCTLVCPDLHNOVTAEDGTQRCCKSPCARGTHSL 345

```

Query Match 2.1%; Score 9; DB 11; Length 478; Best local Similarity 100.0%; Pred. No. 2.4; Mismatches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 239 OCAAGCTGP 247
 DB 235 OCAAGCTGP 243

RESULT 5

O90836 PRELIMINARY: PRT: 527 AA.
 AC O90836:
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE EGF/TCF-alpha receptor precursor.
 GN C-ERBB.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauaria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.

RA MEDLINE=92123214; PubMed=1732751;
 RA Flickinger T.W., Maible N.J., Kung H.-J.;
 RT "An alternatively processed mRNA from the avian c-erbB gene encodes a
 soluble, truncated form of the receptor that can block ligand-
 dependent transfection.";
 RT Mol. Cell. Biol. 12:883-893(1992).

DR EMBL: M77637; AAA48759.1;
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR002174; Furin-like.
 DR Pfam: PF00757; Furin-like; 1.
 DR SMART: SM00261; FUR_2.
 DR SMART: SM00261; FUR_2.

KW Receptor; Signal.
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 527 EGF/TCF-ALPHA RECEPTOR.
 SQ SEQUENCE 527 AA; 58353 MW; 764564ABCC095298 CRC64;

Query Match
 Best Local Similarity 2.1%; Score 9; DB 13; Length 527;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 239 OCAAGCTGP 247
 DB 245 OCAAGCTGP 253

RESULT 6

O9PSH2 PRELIMINARY: PRT: 599 AA.
 AC O9PSH2:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Epidermal growth factor receptor (Fragment).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauaria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.

RA MEDLINE=94020816; PubMed=8414496;
 RA Callaghan T., Antczak M., Flickinger T., Raines M., Myers M.,
 RA Kung H.-J.;
 RT "A complete description of the EGF-receptor exon structure;
 RT implication in oncogenic activation and domain evolution.";
 RT Oncogene 8:2939-2948(1993).

DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR002174; Furin-like.
 DR Pfam: PF00757; Furin-like; 2.
 DR Pfam: PF01030; Recep_L_domain; 2.

DR SMART: SM00261; FUR_3.
 SQ SEQUENCE 599 AA; 66363 MW; FEAB46D23D991BD CRC64;

Query Match
 Best Local Similarity 2.1%; Score 9; DB 13; Length 599;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 239 OCAAGCTGP 247
 DB 393 OCAAGCTGP 401

RESULT 7

O9ERV6 PRELIMINARY: PRT: 643 AA.
 AC O9ERV6:
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Epidermal growth factor receptor isoform 2.
 GN EGFR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=C3H/101, 129/SVJ, AND 129/SVEV/TAC;
 RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,
 RA Sinclair C.S., Pearshall R.S., Green P.J., Yee D., Lampland A.L.,
 RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,
 RA Maible N.J.;

RT "Comparative genomic sequence analysis and isolation of human and
 RT mouse alternative Egr transcripts encoding truncated receptor
 RT isoforms.";
 RT Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.

DR EMBL: AF275366; AAC28046.1; JOINED.
 DR EMBL: AF275364; AAC28046.1; JOINED.
 DR EMBL: AF275365; AAC28046.1; JOINED.
 DR MGD: MGI:95294; Egrf.
 DR InterPro: IPR000345; Cytochrome_bld.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR002174; Furin-like.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR SMART: SM00261; FUR_4.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 KW Receptor.
 SQ SEQUENCE 643 AA; 71476 MW; DEF22002C8491B1 CRC64;

Query Match
 Best Local Similarity 2.1%; Score 9; DB 11; Length 643;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 239 OCAAGCTGP 247
 DB 235 OCAAGCTGP 243

RESULT 8

O9WVF5 PRELIMINARY: PRT: 655 AA.
 AC O9WVF5:
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Epidermal growth factor receptor (Epidermal growth factor receptor
 DE isoform 3).
 GN EGFR.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=LIVER;
RA Reiter J.L., Threaddgill D.W., Danielsen A.J., Schell C.,
RA Lampland A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R.,
RA Maible N.J.;
RT "Alternative transcripts from the Human and Mouse EGFR Genes Encode
RT Carboxy-Terminal Truncated Receptors.";
RN Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H/101, 129/SVJ, AND 129/SVEV7AC;
RA Reiter J.L., Threaddgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,
RA Sinclair C.S., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,
RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,
RA Maible N.J.;
RT "Comparative genomic sequence analysis and isolation of human and
RT mouse alternative Egfr transcripts encoding truncated receptor
RT isoforms.";
RN Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=LIVER;
RX MEDLINE=21083660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Saio T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
RA Kuehl S., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Norioka P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN RT
RT Nature 409:685-690(2001).
DR EMBL: AF124513; AAD4149.1; -
DR EMBL: AF275366; AAG28047.1; -
DR EMBL: AF275364; AAG28047.1; JOINED.
DR EMBL: AF275365; AAG28047.1; JOINED.
DR EMBL: AK004944; BAB23688.1; -
DR EMBL: AK004883; BAB23641.1; -
DR EMBL: AK004911; BAB23662.1; -
DR MGI: 95294; Egfr.
DR InterPro: IPR000494; EGFR_L.domain.
DR InterPro: IPR002174; Furin-like.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF01030; Recep_L.domain; 2.
DR SMART: SM00261; FU; 3.
KW Receptor.
SQ SEQUENCE 655 AA; 72906 MW; 6B34063B1BC928CB CRC64;

```

```

DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Epidermal growth factor receptor.
GN EGFR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FISHER; TISSUE=LIVER;
RX MEDLINE=90258888; PubMed=2342466;
RA Petch L.A., Harris J., Raymond V.W., Blasband A.J., Lee D.C.,
RA Earp H.S.;
RT "A truncated, secreted form of the epidermal growth factor receptor is
RT encoded by an alternatively spliced transcript in normal rat tissue.";
RN Mol. Cell. Biol. 10:2973-2982(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FISHER; TISSUE=LIVER;
RA Petch L.A.;
RN Submitted (NOV-1991) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=FISHER; TISSUE=LIVER;
RX Gutridge K., Dawson T.L., Earp H.S.;
RN Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: M37394; AF144008.1; -
DR HSP: P11362; IFGK.
DR InterPro: IPR000494; EGFR_L.domain.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR002145; Tyr_pkinase.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF00069; Pkinase; 1.
DR Pfam: PF01030; Recep_L.domain; 2.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00261; FU; 3
DR SMART: SM00219; TYRK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; Receptor; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 1209 AA; 134891 MW; 96FE76CC1B7773 CRC64;

```

Query Match 2.1%; Score 9; DB 11; Length 1209;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 239 OCAAGCTGP 247
DB 235 OCAAGCTGP 243

```

RESULT 10
Q9EP98 PRELIMINARY; PRT; 1210 AA.
AC Q9EP98;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Epidermal growth factor receptor isoform 1.
GN EGFR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H/101, 129/SVJ, AND 129/SVEV7AC;
RA Reiter J.L., Threaddgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,


```

RA Sinclair C.S., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,
RA Balasubramanian S., Crossley T.O., Magnuson T.R., James C.D.,
RA Maine N.J.;
RT "Comparative genomic sequence analysis and isolation of human and
RT mouse alternative Egr transcripts encoding truncated receptor
RT isoforms.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,
RA Scheel C., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,
RA Balasubramanian S., Crossley T.O., Magnuson T.R., James C.D.,
RA Maine N.J.;
RT "Comparative genomic sequence analysis and isolation of human and
RT mouse alternative Egr transcripts encoding truncated receptor
RT isoforms.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,
RA Scheel C., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,
RA Balasubramanian S., Crossley T.O., Magnuson T.R., James C.D.,
RA Maine N.J.;
RT "Comparative genomic sequence analysis and isolation of human and
RT mouse alternative Egr transcripts encoding truncated receptor
RT isoforms.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF275366; AAG28045.1; -.
DR EMBL: AF275364; AAG28045.1; JOINED.
DR EMBL: AF275365; AAG28045.1; JOINED.
DR EMBL: AF275367; AAG24386.1; -.
DR HSSP: P11362; 1FGK.
DR MGD: MGI:95294; Egfr.
DR InterPro: IPR000345; Cyto_heme_bind.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00057; Kinase_1.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PD000001; Euk_pkinase; 1.
DR SMART: SM00261; FU; 5.
DR SMART: SM00220; S_TKC; 1.
DR SMART: SM00219; TYRK; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE: PS00107; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR ATP-binding; Receptor; transferase.
SQ SEQUENCE 1210 AA; 134840 MW; 62CD021C9DE32E18 CRC64;

Query Match
Best Local Similarity 100.0%; Score 9; DB 11; Length 1210;
Pred. No. 5.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 239 QCAAGCTGP 247
DB 235 QCAAGCTGP 243

```

```

RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=91172172; PubMed=1672439;
RA Van D.H., Hung M.C.;
RT "Identification and characterization of a novel enhancer for the rat
RT neu promoter.";
RL Mol. Cell. Biol. 11:1875-1882(1991).
DR EMBL: M61004; AAA41686.1; -.
KW Oncogene; Transforming protein.
FT NON_TER
SQ SEQUENCE 48 AA; 4966 MW; 72A0E4B1E13F1129 CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 11; Length 48;
Pred. No. 3.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 LALLPPG 19
DB 15 LALLPPG 22

RESULT 12
O90216 PRELIMINARY; PRT; 63 AA.
AC O90216;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE Carbonyl reductase isoform I (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR; TISSUE=OVARY;
RA Espey L.L., Yoshioka S., Russell D., Ujioka T., Vladu B., Skelsey M.,
RA Fujii S., Okamura H., Richards J.S.;
RT "Characterization of Ovarian Carbonyl Reductase Gene Expression during
RT Ovulation in the Gonadotropin-Primed Immature Rat.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF181955; AAF03394.1; -.
FT NON_TER
SQ SEQUENCE 63 AA; 6947 MW; 03514BC1F399AA56 CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 11; Length 63;
Pred. No. 4.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 LALLPPGA 20
DB 40 LALLPPGA 47

RESULT 13
O90215 PRELIMINARY; PRT; 63 AA.
AC O90215;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE Carbonyl reductase isoform II (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR; TISSUE=OVARY;
RA Espey L.L., Yoshioka S., Russell D., Ujioka T., Vladu B., Skelsey M.,
RA Fujii S., Okamura H., Richards J.S.;
RT "Characterization of Ovarian Carbonyl Reductase Gene Expression during
RT Ovulation in the Gonadotropin-Primed Immature Rat.";

```

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF181956; AAF03395.1; -
 FT NON_TER 1 1
 SQ SEQUENCE 63 AA; 6916 MW; D5964BC1E299A2F4 CRC64;

Query Match 1.9%; Score 8; DB 11; Length 63;
 Best Local Similarity 100.0%; Pred. No. 4.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 LALLPGA 20
 |||||||
 Db 40 LALLPGA 47

RESULT 14

O9N6W6 PRELIMINARY; PRT; 217 AA.

AC O9N6W6; PRELIMINARY; PRT; 217 AA.
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE HLHgamma protein.
 GN HLH-GAMMA OR E.
 OS Drosophila hydei (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7224;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Cooper M.T.D., Tyler D.M., Furiols M., Chalkiadaki A., Delidakis C.,
 RA Bray S.;
 RT "Spatially restricted factors co-operate with Notch in the regulation
 of Enhancer of split genes.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ276313; CAB77019.1;
 DR Flybase; FBgn0012370; Dhyd\HLH-gamma.
 DR InterPro; IPR001092; HLH_basic.
 DR InterPro; IPR003650; Orange.
 DR Pfam; PF00010; HLH; 1.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00511; ORANGE; 1.
 DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
 SQ SEQUENCE 217 AA; 24082 MW; 39177C58F322B4ED CRC64;

Query Match 1.9%; Score 8; DB 5; Length 217;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 392 SVRISPV 399
 |||||||
 Db 159 SVRISPV 166

RESULT 15

O08558 PRELIMINARY; PRT; 277 AA.

AC O08558; PRELIMINARY; PRT; 277 AA.
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Inducible carbonyl reductase.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-MISTER; TISSUE-OVARY;
 RX MEDLINE=97167735; PubMed=9015353;
 RA Aoki H., Okada T., Mizutani T., Numata Y., Minegishi T., Miyamoto K.;
 RT "Identification of two closely related genes, inducible and
 noninducible carbonyl reductases in the rat ovary.";

RL Biochem. Biophys. Res. Commun. 230:518-523(1997).
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 (SDR) FAMILY.
 DR EMBL; D89069; BAAL9007.1; -
 DR HSSP; P50162; 1AE1.
 DR InterPro; IPR002198; ADH_short.
 DR Pfam; PF00106; adh_short; 1.
 DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
 KW Oxidoreductase.

SQ SEQUENCE 277 AA; 30654 MW; DDA015D1B71673A7 CRC64;

Query Match 1.9%; Score 8; DB 11; Length 277;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 LALLPGA 20
 |||||||
 Db 254 LALLPGA 261

Search completed: March 4, 2003, 12:46:40
 Job time : 48.2751 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:43:52 ; Search time 33.6546 Seconds
(without alignments)
366.315 Million cell updates/sec

Title: US-09-234-208b-2

Sequence: 1 MELAALCRWGLLALLPPGA.....VGRGPDPAHVAVNLRYEG 419

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents_AA:*

- 1: /cgn2_6/prodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2_6/prodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2_6/prodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2_6/prodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2_6/prodata/1/1aa/PCU05_COMB.pep.*
- 6: /cgn2_6/prodata/1/1aa/Backfillset1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	419	100.0	419	4	US-09-630-155-2	Sequence 2, Appl1
2	340	81.1	782	2	US-09-146-283-4	Sequence 4, Appl1
3	340	81.1	782	3	US-08-579-823A-4	Sequence 4, Appl1
4	340	81.1	782	4	US-09-344-195-4	Sequence 4, Appl1
5	340	81.1	1255	1	US-08-467-083-68	Sequence 68, Appl1
6	340	81.1	1255	1	US-08-414-417B-68	Sequence 68, Appl1
7	340	81.1	1255	2	US-08-484-438-8	Sequence 8, Appl1
8	340	81.1	1255	2	US-08-486-348A-68	Sequence 8, Appl1
9	340	81.1	1255	2	US-08-625-101-2	Sequence 2, Appl1
10	340	81.1	1255	2	US-08-468-545B-68	Sequence 68, Appl1
11	340	81.1	1255	3	US-08-356-786-2	Sequence 2, Appl1
12	340	81.1	1255	3	US-08-466-680B-68	Sequence 68, Appl1
13	319	76.1	624	3	US-08-422-108-1	Sequence 68, Appl1
14	319	76.1	624	4	US-08-422-734-1	Sequence 1, Appl1
15	96	22.9	97	4	US-08-421-356-3	Sequence 3, Appl1
16	96	22.9	97	4	US-09-046-783-3	Sequence 3, Appl1
17	79	18.9	79	4	US-09-630-155-1	Sequence 1, Appl1
18	17	4.1	17	1	US-08-467-083-61	Sequence 61, Appl1
19	17	4.1	17	1	US-08-414-417B-61	Sequence 61, Appl1
20	17	4.1	17	2	US-08-486-348A-61	Sequence 61, Appl1
21	17	4.1	17	2	US-08-468-545B-61	Sequence 61, Appl1
22	17	4.1	17	3	US-08-466-680B-61	Sequence 61, Appl1
23	15	3.6	15	1	US-08-467-083-30	Sequence 30, Appl1
24	15	3.6	15	1	US-08-467-083-31	Sequence 31, Appl1
25	15	3.6	15	1	US-08-467-083-32	Sequence 32, Appl1
26	15	3.6	15	1	US-08-467-083-33	Sequence 33, Appl1
27	15	3.6	15	1	US-08-467-083-56	Sequence 56, Appl1

28	15	3.6	15	1	US-08-414-417B-30	Sequence 30, Appl1
29	15	3.6	15	1	US-08-414-417B-31	Sequence 31, Appl1
30	15	3.6	15	1	US-08-414-417B-32	Sequence 32, Appl1
31	15	3.6	15	1	US-08-414-417B-33	Sequence 33, Appl1
32	15	3.6	15	1	US-08-414-417B-56	Sequence 56, Appl1
33	15	3.6	15	2	US-08-486-348A-30	Sequence 30, Appl1
34	15	3.6	15	2	US-08-486-348A-31	Sequence 31, Appl1
35	15	3.6	15	2	US-08-486-348A-32	Sequence 32, Appl1
36	15	3.6	15	2	US-08-486-348A-33	Sequence 33, Appl1
37	15	3.6	15	2	US-08-486-348A-56	Sequence 56, Appl1
38	15	3.6	15	2	US-08-468-545B-30	Sequence 30, Appl1
39	15	3.6	15	2	US-08-468-545B-31	Sequence 31, Appl1
40	15	3.6	15	2	US-08-468-545B-32	Sequence 32, Appl1
41	15	3.6	15	2	US-08-468-545B-33	Sequence 33, Appl1
42	15	3.6	15	2	US-08-468-545B-56	Sequence 56, Appl1
43	15	3.6	15	3	US-08-466-680B-30	Sequence 30, Appl1
44	15	3.6	15	3	US-08-466-680B-31	Sequence 31, Appl1
45	15	3.6	15	3	US-08-466-680B-32	Sequence 32, Appl1

ALIGNMENTS

RESULT 1
US-09-630-155-2
; Sequence 2, Application US/09630155
; Patent No. 6414130
GENERAL INFORMATION:
; APPLICANT: Doherty, Joni Kristin and Gail M. Clinton
; TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVIS WRIGHT TREMAINE LLP
; STREET: 1501 Fourth Avenue, 2600 Century Square
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98101
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: PC compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: Word
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/630.155
; FILING DATE: 16-Jan-2001
; CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
; NAME: Davison, Barry L.
; REGISTRATION NUMBER: 47,309
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206 628-7621
; TELEFAX: 206 628-7699
INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: polypeptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-630-155-2
Query Match 100.0%; Score 419; DB 4; Length 419;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MELAALCRWGLLALLPPGAASVQCTGTDMKRLRPASPTHLDMRLHYLGGCVQVGNL 60
|||||
Db 1 MELAALCRWGLLALLPPGAASVQCTGTDMKRLRPASPTHLDMRLHYLGGCVQVGNL 60
QY 61 ELTYLPTNASTLSPDIOEVQGVVLIANHNOVROVPLQRLRIYVGTQLFEDNVALLAVLDNG 120

```
Db 61 ELTYLPTNLSLFLQDIOEVOGVYLLAHNQVRVPLQRLRIYVGTQLFEDNYALAVLDNG 120
Qy 121 DPLNNTPTVGTASPGGLRELQRLSLTEILKGVLIQRPOLCYQDTILMKDIFHKNNOLA 180
Db 121 DPLNNTPTVGTASPGGLRELQRLSLTEILKGVLIQRPOLCYQDTILMKDIFHKNNOLA 180
Qy 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDDCSLRTVACAGCARKGRLPTDCCHEOC 240
Db 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDDCSLRTVACAGCARKGRLPTDCCHEOC 240
Qy 241 AACGTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYTACP 300
Db 241 AACGTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYTACP 300
Qy 301 YNLTSTDVGSCTLVCPHLNQEVTAEQDGTORCEKSKPCARGTSHLLPRPAVVPPLRMOP 360
Db 301 YNLTSTDVGSCTLVCPHLNQEVTAEQDGTORCEKSKPCARGTSHLLPRPAVVPPLRMOP 360
Qy 361 GPAHPVLSFLRSPMDLVSAFYSPLAPLSPVSPISVGRGPPDDAHAVANLSRYEG 419
Db 361 GPAHPVLSFLRSPMDLVSAFYSPLAPLSPVSPISVGRGPPDDAHAVANLSRYEG 419
```

RESULT 2

```
US-09-146-283-4
; Sequence 4, Application US/09146283
; Patent No. 5976546
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 782 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
US-09-146-283-4
```

```
Query Match 81.1%; Score 340; DB 2; Length 782;
Best Local Similarity 100.0%; Pred. No. 9.7e-317;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 METALCRWGLLALLPPAASSTOVCTGDMKRLPASPETHLMDLRHLXYGCGVYQGNL 60
Db 1 METALCRWGLLALLPPAASSTOVCTGDMKRLPASPETHLMDLRHLXYGCGVYQGNL 60
Qy 61 ELTYLPTNLSLFLQDIOEVOGVYLLAHNQVRVPLQRLRIYVGTQLFEDNYALAVLDNG 120
Db 61 ELTYLPTNLSLFLQDIOEVOGVYLLAHNQVRVPLQRLRIYVGTQLFEDNYALAVLDNG 120
Qy 121 DPLNNTPTVGTASPGGLRELQRLSLTEILKGVLIQRPOLCYQDTILMKDIFHKNNOLA 180
Db 121 DPLNNTPTVGTASPGGLRELQRLSLTEILKGVLIQRPOLCYQDTILMKDIFHKNNOLA 180
Qy 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDDCSLRTVACAGCARKGRLPTDCCHEOC 240
Db 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDDCSLRTVACAGCARKGRLPTDCCHEOC 240
Qy 241 AACGTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYTACP 300
Db 241 AACGTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYTACP 300
Qy 301 YNLTSTDVGSCTLVCPHLNQEVTAEQDGTORCEKSKPCAR 340
Db 301 YNLTSTDVGSCTLVCPHLNQEVTAEQDGTORCEKSKPCAR 340
```

RESULT 3

```
US-08-579-823A-4
; Sequence 4, Application US/08579823A
; Patent No. 6080409
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Composition and Method
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/579,823A
; FILING DATE: 03-DEC-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 782 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
US-08-579-823A-4
```

```
Query Match 81.1%; Score 340; DB 3; Length 782;
Best Local Similarity 100.0%; Pred. No. 9.7e-317;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MEALALCRWGLLALLPPGAASVQVCTGTDMLRLPASPTHTDMLRHLVQGGQVQVGNL 60
Db 1 MEALALCRWGLLALLPPGAASVQVCTGTDMLRLPASPTHTDMLRHLVQGGQVQVGNL 60
QY 61 ELTYLPTNASTLFDIOIEVQGVYLLAHNOVROPYLRRLIVGTQOLFEDNVALAVLDNG 120
Db 61 ELTYLPTNASTLFDIOIEVQGVYLLAHNOVROPYLRRLIVGTQOLFEDNVALAVLDNG 120
QY 121 DPLNNTPTVYGASPGGLRELQLSLFEILKGVLIORNPOLCYODITLMKDFHKNQOLA 180
Db 121 DPLNNTPTVYGASPGGLRELQLSLFEILKGVLIORNPOLCYODITLMKDFHKNQOLA 180
QY 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDCOSLRTVCAGGCARCKGFLPTDCHEQC 240
Db 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDCOSLRTVCAGGCARCKGFLPTDCHEQC 240
QY 241 AACCTGPKHSDCLACHFNHSGICEHLCPALVYNTDTEFSMPNPGRYTFGASCTYACP 300
Db 241 AACCTGPKHSDCLACHFNHSGICEHLCPALVYNTDTEFSMPNPGRYTFGASCTYACP 300
QY 301 YNYLSTDVGSCTLVCPPLHNOEVTAEADGTORCEKSKPCAR 340
Db 301 YNYLSTDVGSCTLVCPPLHNOEVTAEADGTORCEKSKPCAR 340
```

RESULT 4

US-09-344-195-4

; Sequence 4, Application US/09344195
; Patent No. 6210662

GENERAL INFORMATION:

APPLICANT: Laus, Retner
Ruegg, Curtis L.

Mu, Hongyu

TITLE OF INVENTION: Immunostimulatory Compositions

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSER: Dehlinger & Associates

STREET: 350 Cambridge Ave. Suite 250

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/344,195

FILING DATE: 24-Jun-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/146,283

FILING DATE: 03-SEPT-1998

ATTORNEY/AGENT INFORMATION:

NAME: Judge, Linda R.

REGISTRATION NUMBER: 42,702

REFERENCE/DOCKET NUMBER: 7636-0010.21

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-324-0880

TELEFAX: 650-324-0960

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 782 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

ORGANISM: homo sapiens

INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-344-195-4

Query Match 81.1%; Score 340; DB 4; Length 782;

Best Local Similarity 100.0%; Pred No. 9.7e-317;

Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MEALALCRWGLLALLPPGAASVQVCTGTDMLRLPASPTHTDMLRHLVQGGQVQVGNL 60
Db 1 MEALALCRWGLLALLPPGAASVQVCTGTDMLRLPASPTHTDMLRHLVQGGQVQVGNL 60
QY 61 ELTYLPTNASTLFDIOIEVQGVYLLAHNOVROPYLRRLIVGTQOLFEDNVALAVLDNG 120
Db 61 ELTYLPTNASTLFDIOIEVQGVYLLAHNOVROPYLRRLIVGTQOLFEDNVALAVLDNG 120
QY 121 DPLNNTPTVYGASPGGLRELQLSLFEILKGVLIORNPOLCYODITLMKDFHKNQOLA 180
Db 121 DPLNNTPTVYGASPGGLRELQLSLFEILKGVLIORNPOLCYODITLMKDFHKNQOLA 180
QY 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDCOSLRTVCAGGCARCKGFLPTDCHEQC 240
Db 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDCOSLRTVCAGGCARCKGFLPTDCHEQC 240
QY 241 AACCTGPKHSDCLACHFNHSGICEHLCPALVYNTDTEFSMPNPGRYTFGASCTYACP 300
Db 241 AACCTGPKHSDCLACHFNHSGICEHLCPALVYNTDTEFSMPNPGRYTFGASCTYACP 300
QY 301 YNYLSTDVGSCTLVCPPLHNOEVTAEADGTORCEKSKPCAR 340
Db 301 YNYLSTDVGSCTLVCPPLHNOEVTAEADGTORCEKSKPCAR 340
```

RESULT 5

US-08-467-083-68

; Sequence 68, Application US/08467083
; Patent No. 5726023

GENERAL INFORMATION:

APPLICANT: Cheever, Martin A.

APPLICANT: Disis, Mary L.

TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN

TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE

NUMBER OF SEQUENCES: 68

CORRESPONDENCE ADDRESS:

ADDRESSER: Seed and Berry

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: US

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467,083

FILING DATE: 06-JUN-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/414,417

FILING DATE: 06-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Sharkey, Richard G.

REGISTRATION NUMBER: 32,629

REFERENCE/DOCKET NUMBER: 920010.448C2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

TELEX: 3723836 SEEDANBERY

INFORMATION FOR SEQ ID NO: 68:

SEQUENCE CHARACTERISTICS:

LENGTH: 1255 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-467-083-68

Query Match 81.1%; Score 340; DB 1; Length 1255;

Best Local Similarity 100.0%; Pred. No. 1.5e-316; Mismatches 0; Indels 0; Gaps 0;

Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELAALCRWGLLALLPPGASQVCTGTDMLRLPASFETHDMLRHLXYGCGVQVGNL 60
DB 1 MELAALCRWGLLALLPPGASQVCTGTDMLRLPASFETHDMLRHLXYGCGVQVGNL 60
QY 61 ELTYLPNLSLFLDIOEYGVYLAHNOVROPLORLRYRGTDLPEDNVALAVLDNG 120
DB 61 ELTYLPNLSLFLDIOEYGVYLAHNOVROPLORLRYRGTDLPEDNVALAVLDNG 120
QY 121 DPLNNTPTVGTGSPGGLRELQRLSLTEILKGVLIORNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVGTGSPGGLRELQRLSLTEILKGVLIORNPOLCYODTILMKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDCOSLTRVCAGGACARCKGPLPDCHEQC 240
DB 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDCOSLTRVCAGGACARCKGPLPDCHEQC 240
QY 241 AACCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMPNPGRTTFGASCTYACP 300
DB 241 AACCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMPNPGRTTFGASCTYACP 300
QY 301 YNYLSTDVGSCTLVCPHLHNOEVTAEEDGTORCEKCSKPCAR 340
DB 301 YNYLSTDVGSCTLVCPHLHNOEVTAEEDGTORCEKCSKPCAR 340

RESULT 6

US-08-414-417B-68

Sequence 68, Application US/08414417B

Patent No. 5801005

GENERAL INFORMATION:

APPLICANT: Cheever, Martin A.

APPLICANT: Disis, Mary L.

TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN

TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE

TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED

NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Seed and Berry LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: US

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/414,417B

FILING DATE: 31-MAR-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Sharkey, Richard G.

REGISTRATION NUMBER: 32,629

REFERENCE/DOCKET NUMBER: 920010.448C2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 68:

SEQUENCE CHARACTERISTICS:

LENGTH: 1255 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-414-417B-68

Query Match 81.1%; Score 340; DB 1; Length 1255;

Best Local Similarity 100.0%; Pred. No. 1.5e-316; Mismatches 0; Indels 0; Gaps 0;

Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELAALCRWGLLALLPPGASQVCTGTDMLRLPASFETHDMLRHLXYGCGVQVGNL 60
DB 1 MELAALCRWGLLALLPPGASQVCTGTDMLRLPASFETHDMLRHLXYGCGVQVGNL 60
QY 61 ELTYLPNLSLFLDIOEYGVYLAHNOVROPLORLRYRGTDLPEDNVALAVLDNG 120
DB 61 ELTYLPNLSLFLDIOEYGVYLAHNOVROPLORLRYRGTDLPEDNVALAVLDNG 120
QY 121 DPLNNTPTVGTGSPGGLRELQRLSLTEILKGVLIORNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVGTGSPGGLRELQRLSLTEILKGVLIORNPOLCYODTILMKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDCOSLTRVCAGGACARCKGPLPDCHEQC 240
DB 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDCOSLTRVCAGGACARCKGPLPDCHEQC 240
QY 241 AACCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMPNPGRTTFGASCTYACP 300
DB 241 AACCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMPNPGRTTFGASCTYACP 300
QY 301 YNYLSTDVGSCTLVCPHLHNOEVTAEEDGTORCEKCSKPCAR 340
DB 301 YNYLSTDVGSCTLVCPHLHNOEVTAEEDGTORCEKCSKPCAR 340

RESULT 7

US-08-484-438-8

Sequence 8, Application US/08484438

Patent No. 5811098

Patent No. 5811098 5780031

GENERAL INFORMATION:

APPLICANT: Plovman, Gregory D.

APPLICANT: Culouscou, Jean-Michel

APPLICANT: Shoyab, Mohammed

APPLICANT: Siegal, Clay B.

APPLICANT: Helster m, Ingegerd

TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE

TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,438

FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/323,442

FILING DATE: 14-OCT-1994

APPLICATION NUMBER: US 08/150,704

FILING DATE: 10-NOV-1993

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/981,165

FILING DATE: 24-NOV-1992

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Mistic, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 5624-230

```
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-484-438-8

Query Match
Best local Similarity 100.0%; Pred. No. 1.5e-316; Length 1255;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELAALCRWGLLALLPFGAASFOVCTGDMKRLPASPEHLDMLRHLYQGCQVQGNL 60
DB 1 MELAALCRWGLLALLPFGAASFOVCTGDMKRLPASPEHLDMLRHLYQGCQVQGNL 60
QY 61 ELTYLPTNASLSFLQDIEVQGVVLAHNOVROVPLQRLRIYVGTQLFEDNVALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIEVQGVVLAHNOVROVPLQRLRIYVGTQLFEDNVALAVLDNG 120
QY 121 DPLNNTPTVYGASPGGLRELQRLSLTEILKGVLIQRPOLCYQDPTLMKDIFHKNNOLA 180
DB 121 DPLNNTPTVYGASPGGLRELQRLSLTEILKGVLIQRPOLCYQDPTLMKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMCKSGRCWSESSDCSLTRTVAGGACARCKGPLPDDCHEQC 240
DB 181 LTLIDNRSRACHPCSPMCKSGRCWSESSDCSLTRTVAGGACARCKGPLPDDCHEQC 240
QY 241 AACGTGPKHSDCLACHFNHSGICELHCPALVYNTDFESMNPBGRYFFGASCVYACP 300
DB 241 AACGTGPKHSDCLACHFNHSGICELHCPALVYNTDFESMNPBGRYFFGASCVYACP 300
QY 301 YNYLSTDVGSCTLVCPHLNQEVTAEEDGTORCEKSRPCAR 340
DB 301 YNYLSTDVGSCTLVCPHLNQEVTAEEDGTORCEKSRPCAR 340

RESULT 8
US-08-486-348A-68
; Sequence 68, Application US/08486348A
; Patent No. 5846538
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; TITLE OF INVENTION: DISIS, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,348A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C6
```

```
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-486-348A-68

Query Match
Best local Similarity 100.0%; Pred. No. 1.5e-316; Length 1255;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELAALCRWGLLALLPFGAASFOVCTGDMKRLPASPEHLDMLRHLYQGCQVQGNL 60
DB 1 MELAALCRWGLLALLPFGAASFOVCTGDMKRLPASPEHLDMLRHLYQGCQVQGNL 60
QY 61 ELTYLPTNASLSFLQDIEVQGVVLAHNOVROVPLQRLRIYVGTQLFEDNVALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIEVQGVVLAHNOVROVPLQRLRIYVGTQLFEDNVALAVLDNG 120
QY 121 DPLNNTPTVYGASPGGLRELQRLSLTEILKGVLIQRPOLCYQDPTLMKDIFHKNNOLA 180
DB 121 DPLNNTPTVYGASPGGLRELQRLSLTEILKGVLIQRPOLCYQDPTLMKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMCKSGRCWSESSDCSLTRTVAGGACARCKGPLPDDCHEQC 240
DB 181 LTLIDNRSRACHPCSPMCKSGRCWSESSDCSLTRTVAGGACARCKGPLPDDCHEQC 240
QY 241 AACGTGPKHSDCLACHFNHSGICELHCPALVYNTDFESMNPBGRYFFGASCVYACP 300
DB 241 AACGTGPKHSDCLACHFNHSGICELHCPALVYNTDFESMNPBGRYFFGASCVYACP 300
QY 301 YNYLSTDVGSCTLVCPHLNQEVTAEEDGTORCEKSRPCAR 340
DB 301 YNYLSTDVGSCTLVCPHLNQEVTAEEDGTORCEKSRPCAR 340

RESULT 9
US-08-625-101-2
; Sequence 2, Application US/08625101
; Patent No. 5869445
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
; TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
; TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
; TITLE OF INVENTION: ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/625,101
; FILING DATE: 01-APR-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
```

TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-625-101-2

Query Match 81.1%; Score 340; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.5e-316;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELAALCRWGLLALLPPGAASVQVCTGTDMLRLPASPEHIDMLRHLYQGCQVQVGNL 60
DB 1 MELAALCRWGLLALLPPGAASVQVCTGTDMLRLPASPEHIDMLRHLYQGCQVQVGNL 60
QY 61 ELTYLPTNASLSFLQDIOEVQGVYLAHNOVROVPLQRLRIVGTQLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIOEVQGVYLAHNOVROVPLQRLRIVGTQLFEDNYALAVLDNG 120
QY 121 DPLNNTPTVYGASPGGLRELQRLSLTEILKGVLIQRPQLCYQDTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVYGASPGGLRELQRLSLTEILKGVLIQRPQLCYQDTILMKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMCKGSRCKWGESSEDCQSLRTVCAGGCARCKGPLPTDCHEQC 240
DB 181 LTLIDNRSRACHPCSPMCKGSRCKWGESSEDCQSLRTVCAGGCARCKGPLPTDCHEQC 240
QY 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCYTACP 300
DB 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCYTACP 300
QY 301 YNLTSDVGSCTLVCPLHNOEVAEDGTORCEKCSKPCAR 340
DB 301 YNLTSDVGSCTLVCPLHNOEVAEDGTORCEKCSKPCAR 340

RESULT 10

US-08-468-545B-68
Sequence 68, Application US/08468545B
Patent No. 5876712
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Disis, Mary L.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,545B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Shaik, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:

SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-468-545B-68

Query Match 81.1%; Score 340; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.5e-316;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELAALCRWGLLALLPPGAASVQVCTGTDMLRLPASPEHIDMLRHLYQGCQVQVGNL 60
DB 1 MELAALCRWGLLALLPPGAASVQVCTGTDMLRLPASPEHIDMLRHLYQGCQVQVGNL 60
QY 61 ELTYLPTNASLSFLQDIOEVQGVYLAHNOVROVPLQRLRIVGTQLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIOEVQGVYLAHNOVROVPLQRLRIVGTQLFEDNYALAVLDNG 120
QY 121 DPLNNTPTVYGASPGGLRELQRLSLTEILKGVLIQRPQLCYQDTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVYGASPGGLRELQRLSLTEILKGVLIQRPQLCYQDTILMKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMCKGSRCKWGESSEDCQSLRTVCAGGCARCKGPLPTDCHEQC 240
DB 181 LTLIDNRSRACHPCSPMCKGSRCKWGESSEDCQSLRTVCAGGCARCKGPLPTDCHEQC 240
QY 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCYTACP 300
DB 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCYTACP 300
QY 301 YNLTSDVGSCTLVCPLHNOEVAEDGTORCEKCSKPCAR 340
DB 301 YNLTSDVGSCTLVCPLHNOEVAEDGTORCEKCSKPCAR 340

RESULT 11

US-08-356-786-2
Sequence 2, Application US/08356786
Patent No. 5877305
GENERAL INFORMATION:
APPLICANT: Huston, James S.
APPLICANT: Oppermann, Hermann
APPLICANT: Houston, L. L.
APPLICANT: Ring, David B.
TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
TITLE OF INVENTION: Marker
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,786
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/831,967
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: CRP-053
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-356-786-2

Query Match 81.1%; Score 340; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.5e-316;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELAALCRWGLLALLPPGAASVCTGTDMKRLRPASPEHLDMLRHLYGCGQVVOGNTL 60
DB 1 MELAALCRWGLLALLPPGAASVCTGTDMKRLRPASPEHLDMLRHLYGCGQVVOGNTL 60
QY 61 ELTYLPTNLSLFDIOIOVGQVLIANHNOVQVPLQRLRIYRGTOLEFDNTALAVLDNG 120
DB 61 ELTYLPTNLSLFDIOIOVGQVLIANHNOVQVPLQRLRIYRGTOLEFDNTALAVLDNG 120
QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQRPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQRPOLCYODTILMKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMKGSRGWESSEDCQSLRTVYACAGGACARCKGPLPTDCHEQC 240
DB 181 LTLIDNRSRACHPCSPMKGSRGWESSEDCQSLRTVYACAGGACARCKGPLPTDCHEQC 240
QY 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCYTACP 300
DB 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCYTACP 300
QY 301 YNVLSTDVGSCTLYCPPLHNOEYTAEDGTORCEKCSKPCAR 340
DB 301 YNVLSTDVGSCTLYCPPLHNOEYTAEDGTORCEKCSKPCAR 340

RESULT 12

US-08-466-680B-68

Sequence 68, Application US/08466680B

Patent No. 6075122

GENERAL INFORMATION:

APPLICANT: Cheever, Martin A.

APPLICANT: Disis, Mary L.

TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN

TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE

NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: US

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PC-DOS/MS-DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,680B

FILING DATE: 06-JUN-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Sharkey, Richard G.

REGISTRATION NUMBER: 32,629

REFERENCE/DOCKET NUMBER: 920010.448C4

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 68:

SEQUENCE CHARACTERISTICS:

LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-466-680B-68

Query Match 81.1%; Score 340; DB 3; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.5e-316;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELAALCRWGLLALLPPGAASVCTGTDMKRLRPASPEHLDMLRHLYGCGQVVOGNTL 60
DB 1 MELAALCRWGLLALLPPGAASVCTGTDMKRLRPASPEHLDMLRHLYGCGQVVOGNTL 60
QY 61 ELTYLPTNLSLFDIOIOVGQVLIANHNOVQVPLQRLRIYRGTOLEFDNTALAVLDNG 120
DB 61 ELTYLPTNLSLFDIOIOVGQVLIANHNOVQVPLQRLRIYRGTOLEFDNTALAVLDNG 120
QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQRPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQRPOLCYODTILMKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMKGSRGWESSEDCQSLRTVYACAGGACARCKGPLPTDCHEQC 240
DB 181 LTLIDNRSRACHPCSPMKGSRGWESSEDCQSLRTVYACAGGACARCKGPLPTDCHEQC 240
QY 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCYTACP 300
DB 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCYTACP 300
QY 301 YNVLSTDVGSCTLYCPPLHNOEYTAEDGTORCEKCSKPCAR 340
DB 301 YNVLSTDVGSCTLYCPPLHNOEYTAEDGTORCEKCSKPCAR 340

RESULT 13

US-08-422-108-1

Sequence 1, Application US/08422108

Patent No. 6015567

GENERAL INFORMATION:

APPLICANT: Hudziak, Robert M.

APPLICANT: Shepard, H. Michael

APPLICANT: Ullrich, Axel

TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Winpatin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/422,108

FILING DATE: 14-Apr-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/355460

FILING DATE: 13-DEC-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/048346

FILING DATE: 15-APR-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/354319

FILING DATE: 19-MAY-1989

ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M

REGISTRATION NUMBER: 00,000

REFERENCE/DOCKET NUMBER: 554C2D2

```
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 624 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-422-108-1

Query Match
Best Local Similarity 76.1%; Score 319; DB 3; Length 624;
Best Local Similarity 100.0%; Pred. No. 9.4e-297;
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 STQVCTGTMKRLPASPETHLDMRLHYQGCVVGNLELYLPTNASTSLFLQDIQEVQ 81
DB 1 STQVCTGTMKRLPASPETHLDMRLHYQGCVVGNLELYLPTNASTSLFLQDIQEVQ 60
QY 82 GYVLIANQVROVPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTPTVYGASPGGLREIQ 141
DB 61 GYVLIANQVROVPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTPTVYGASPGGLREIQ 120
QY 142 LRSLEILKGVLIQIRNPOLCYQDTILMKDIFHKNNQALATLIDNRSRACHPCSPMCKG 201
DB 121 LRSLEILKGVLIQIRNPOLCYQDTILMKDIFHKNNQALATLIDNRSRACHPCSPMCKG 180
QY 202 SRCWGBSSSEDCOSLTRVYAGGACARCKGPLPTDCCHEQCAAGCTGPKHSDCLACHFNHS 261
DB 181 SRCWGBSSSEDCOSLTRVYAGGACARCKGPLPTDCCHEQCAAGCTGPKHSDCLACHFNHS 240
QY 262 GICELHCPALVYNTDPEFSPMPBEGRYTFGASCVYACPYNYLSDVGSCTLVCPHLNQE 321
DB 241 GICELHCPALVYNTDPEFSPMPBEGRYTFGASCVYACPYNYLSDVGSCTLVCPHLNQE 300
QY 322 VTAE DSTQRCCKSKPCAR 340
DB 301 VTAE DSTQRCCKSKPCAR 319

RESULT 14
US-08-422-734-1
Sequence 1, Application US/08422734
Patent No. 6333169
GENERAL INFORMATION:
APPLICANT: Hudziak, Robert M.
APPLICANT: Shepard, H. Michael
APPLICANT: Ullrich, Axel
TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,734
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/422108
FILING DATE: 14-Apr-1995
APPLICATION NUMBER: 08/355460
FILING DATE: 13-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/048346
```

```
FILING DATE: 15-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/354319
FILING DATE: 19-MAY-1989
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 554C2D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 624 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-422-734-1

Query Match
Best Local Similarity 76.1%; Score 319; DB 4; Length 624;
Best Local Similarity 100.0%; Pred. No. 9.4e-297;
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 STQVCTGTMKRLPASPETHLDMRLHYQGCVVGNLELYLPTNASTSLFLQDIQEVQ 81
DB 1 STQVCTGTMKRLPASPETHLDMRLHYQGCVVGNLELYLPTNASTSLFLQDIQEVQ 60
QY 82 GYVLIANQVROVPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTPTVYGASPGGLREIQ 141
DB 61 GYVLIANQVROVPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTPTVYGASPGGLREIQ 120
QY 142 LRSLEILKGVLIQIRNPOLCYQDTILMKDIFHKNNQALATLIDNRSRACHPCSPMCKG 201
DB 121 LRSLEILKGVLIQIRNPOLCYQDTILMKDIFHKNNQALATLIDNRSRACHPCSPMCKG 180
QY 202 SRCWGBSSSEDCOSLTRVYAGGACARCKGPLPTDCCHEQCAAGCTGPKHSDCLACHFNHS 261
DB 181 SRCWGBSSSEDCOSLTRVYAGGACARCKGPLPTDCCHEQCAAGCTGPKHSDCLACHFNHS 240
QY 262 GICELHCPALVYNTDPEFSPMPBEGRYTFGASCVYACPYNYLSDVGSCTLVCPHLNQE 321
DB 241 GICELHCPALVYNTDPEFSPMPBEGRYTFGASCVYACPYNYLSDVGSCTLVCPHLNQE 300
QY 322 VTAE DSTQRCCKSKPCAR 340
DB 301 VTAE DSTQRCCKSKPCAR 319

RESULT 15
US-08-421-356-3
Sequence 3, Application US/08421356
Patent No. 5783404
GENERAL INFORMATION:
APPLICANT: Koski, Raymond A.
APPLICANT: Koski, Raymond A.
TITLE OF INVENTION: HER-2
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 Dehavenland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/421,356
FILING DATE:
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
```

NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-327
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 97 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-421-356-3

Query Match 22.9%; Score 96; DB 1; Length 97;
Best Local Similarity 100.0%; Pred. No. 2.9e-84;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 LQRLRIVRGTOLEEDNYALAVLDNGDPLNNTTPVVGASPGGLRELOLRSLTEILKGVLI 155
|||||
DB 2 LQRLRIVRGTOLEEDNYALAVLDNGDPLNNTTPVVGASPGGLRELOLRSLTEILKGVLI 61
QY 156 QRNPOLCYODTILMKDIFHKNNQLATLIDTNRRA 191
|||||
DB 62 QRNPOLCYODTILMKDIFHKNNQLATLIDTNRRA 97

Search completed: March 4, 2003, 12:48:29
Job time : 35.6546 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Comphen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 12:46:47 ; Search time 17.6687 Seconds
(without alignments)
1000.035 Million cell updates/sec

Title: US-09-234-208B-2

Perfect score: 419

Sequence: 1 METALCRWGLLALLPPGA.....VGRGPDPAHVAVNLRYEG 419

Scoring table: OLIGO

Gapop 60.0 , Gapect 60.0

Searched: 188354 seqs, 42170167 residues

Word size : 0

Total number of hits satisfying chosen parameters: 188354

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published_Applications_AA.*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCITUS_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCITUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	340	81.1	645	10 US-09-921-161-1	Sequence 1, Appli
2	340	81.1	653	9 US-09-854-356-3	Sequence 3, Appli
3	340	81.1	712	9 US-09-854-356-7	Sequence 7, Appli
4	340	81.1	919	9 US-09-854-356-6	Sequence 6, Appli
5	340	81.1	1235	9 US-09-769-508-2	Sequence 2, Appli
6	340	81.1	1255	9 US-09-854-356-1	Sequence 1, Appli
7	340	81.1	1255	9 US-09-930-125-2	Sequence 2, Appli
8	340	81.1	1255	9 US-09-441-411-6	Sequence 6, Appli
9	340	81.1	1255	10 US-09-811-123-9	Sequence 9, Appli
10	340	81.1	1255	10 US-09-811-115-3	Sequence 9, Appli
11	292	69.7	479	10 US-09-821-883-5	Sequence 5, Appli
12	292	69.7	555	10 US-09-821-883-1	Sequence 1, Appli
13	292	69.7	564	10 US-09-821-883-3	Sequence 3, Appli
14	292	69.7	690	10 US-09-821-883-2	Sequence 2, Appli
15	292	69.7	697	10 US-09-821-883-4	Sequence 4, Appli
16	289	69.0	289	10 US-09-821-883-4	Sequence 23, Appli
17	191	45.6	191	9 US-09-441-411-9	Sequence 9, Appli
18	52	12.4	654	9 US-09-854-356-8	Sequence 8, Appli
19	52	12.4	1256	9 US-09-854-356-2	Sequence 2, Appli

20	52	12.4	1260	9 US-09-870-759-118	Sequence 118, App
21	44	10.5	1256	9 US-09-854-356-14	Sequence 14, Appl
22	22	5.3	22	10 US-09-466-320-19	Sequence 19, Appl
23	22	5.3	68	10 US-09-466-320-11	Sequence 11, Appl
24	18	4.3	19	10 US-09-466-320-20	Sequence 20, Appl
25	18	4.3	65	10 US-09-466-320-12	Sequence 12, Appl
26	15	3.6	15	10 US-09-888-721-21	Sequence 21, Appl
27	14	3.3	14	9 US-10-001-546-31	Sequence 31, Appl
28	10	2.4	10	9 US-10-001-546-19	Sequence 30, Appl
29	9	2.1	9	9 US-09-017-743C-70	Sequence 70, Appl
30	9	2.1	9	9 US-10-001-546-17	Sequence 17, Appl
31	31	2.1	9	9 US-10-001-546-20	Sequence 20, Appl
32	32	2.1	9	9 US-10-001-546-21	Sequence 21, Appl
33	33	2.1	9	9 US-10-001-546-22	Sequence 22, Appl
34	34	2.1	478	10 US-09-867-521-2	Sequence 2, Appli
35	35	2.1	1210	10 US-09-725-433-2	Sequence 2, Appli
36	36	1.9	8	9 US-09-017-743C-92	Sequence 92, Appl
37	37	1.9	9	9 US-09-017-743C-105	Sequence 105, App
38	38	1.9	9	9 US-10-001-546-19	Sequence 19, Appl
39	39	1.9	10	9 US-10-001-546-18	Sequence 18, Appl
40	40	1.9	375	10 US-09-815-242-13383	Sequence 13383, A
41	41	1.9	375	10 US-09-815-242-13700	Sequence 13700, A
42	42	1.9	602	10 US-09-815-242-12896	Sequence 12896, A
43	43	1.7	35	9 US-09-981-876-204	Sequence 204, App
44	44	1.7	35	9 US-09-148-545-204	Sequence 204, App
45	45	1.7	36	10 US-09-864-761-48790	Sequence 48790, A

ALIGNMENTS

RESULT 1

US-09-921-161-1
; Sequence 1, Application US/09921161
; Patent No. US2002009062A1
; GENERAL INFORMATION:
; APPLICANT: Ralph, Peter
; TITLE OF INVENTION: ANALYTICAL METHOD
; FILE REFERENCE: GENENT 066A US/09/921,161
; CURRENT APPLICATION NUMBER: 60/225,433
; PRIOR APPLICATION NUMBER: 60/225,433
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 645
; TYPE: PRP
; ORGANISM: Homo sapiens
US-09-921-161-1

Query Match 81.1%; Score 340; DB 10; Length 645;

Best Local Similarity 100.0%; Pred.No. 9.4e-303;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 METALCRWGLLALLPPGAASVOVCTGDMKILRLPASPETHIDMIRHLYOGCQVQGNL 60
|||||
DB 1 METALCRWGLLALLPPGAASVOVCTGDMKILRLPASPETHIDMIRHLYOGCQVQGNL 60
|||||
QY 61 ELTYLPTNLSLFLDIOEVQGVYLAHQVROVPLQRLIRYRGQOLFEDNYALAVLNG 120
|||||
DB 61 ELTYLPTNLSLFLDIOEVQGVYLAHQVROVPLQRLIRYRGQOLFEDNYALAVLNG 120
|||||
QY 121 DPLNNTTPTVGTASPGGLRELQRLSLEILKGVLIQRPOLCYODTILMKDIFHKNNQA 180
|||||
DB 121 DPLNNTTPTVGTASPGGLRELQRLSLEILKGVLIQRPOLCYODTILMKDIFHKNNQA 180
|||||
QY 181 LTLIDNRRACHPCSPMKSGRCWSESESDQSLTRVCAGGACRCKRPLPTDCHEQC 240
|||||
DB 181 LTLIDNRRACHPCSPMKSGRCWSESESDQSLTRVCAGGACRCKRPLPTDCHEQC 240
|||||
QY 241 AAGCTGPKHSDCLACIHFHNSGICELCHPALVTYNTDFESMPNPEGRTYFASCVTACP 300
|||||

Db 241 AAGCTGPKHSDCLACLFHFNHSGICELHCPALVTYNTDFESHPNPEGRTYFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPPLHNOEYTAEDGTORCEKSKPCAR 340
Db 301 YNYLSTDVGSCTLVCPPLHNOEYTAEDGTORCEKSKPCAR 340

RESULT 2
US-09-854-356-3

; Sequence 3, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 3
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
US-09-854-356-3

Query Match 81.1%; Score 340; DB 9; Length 653;
Best Local Similarity 100.0%; Pred. No. 9.5e-303;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELAALCRWGLLLALLPFGASTVOCTGDMKRLRPASPETHLDMRLHLYQGCQVVGNTL 60
Db 1 MELAALCRWGLLLALLPFGASTVOCTGDMKRLRPASPETHLDMRLHLYQGCQVVGNTL 60
QY 61 ELTYLPTNASLSFLDIOGVGVVLAHNOVROVPLQRLRIYRGQLFEDNVALAVLDNG 120
Db 61 ELTYLPTNASLSFLDIOGVGVVLAHNOVROVPLQRLRIYRGQLFEDNVALAVLDNG 120
QY 121 DPLNNTPTVTGASPGGLRLQLRSLEILKGVLIQRPOLCYODTILMKDIFHKNQOLA 180
Db 121 DPLNNTPTVTGASPGGLRLQLRSLEILKGVLIQRPOLCYODTILMKDIFHKNQOLA 180
QY 181 LFLIDTNRSRACHPCSPMKGSRGWSESDCQSLRTVACAGGACARCKGPLPTDCCHQC 240
Db 181 LFLIDTNRSRACHPCSPMKGSRGWSESDCQSLRTVACAGGACARCKGPLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFHFNHSGICELHCPALVTYNTDFESHPNPEGRTYFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHFNHSGICELHCPALVTYNTDFESHPNPEGRTYFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPPLHNOEYTAEDGTORCEKSKPCAR 340
Db 301 YNYLSTDVGSCTLVCPPLHNOEYTAEDGTORCEKSKPCAR 340

RESULT 3

US-09-854-356-7
; Sequence 7, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins

; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 7
; LENGTH: 712
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: fusion protein
US-09-854-356-7

Query Match 81.1%; Score 340; DB 9; Length 712;
Best Local Similarity 100.0%; Pred. No. 1e-302;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELAALCRWGLLLALLPFGASTVOCTGDMKRLRPASPETHLDMRLHLYQGCQVVGNTL 60
Db 1 MELAALCRWGLLLALLPFGASTVOCTGDMKRLRPASPETHLDMRLHLYQGCQVVGNTL 60
QY 61 ELTYLPTNASLSFLDIOGVGVVLAHNOVROVPLQRLRIYRGQLFEDNVALAVLDNG 120
Db 61 ELTYLPTNASLSFLDIOGVGVVLAHNOVROVPLQRLRIYRGQLFEDNVALAVLDNG 120
QY 121 DPLNNTPTVTGASPGGLRLQLRSLEILKGVLIQRPOLCYODTILMKDIFHKNQOLA 180
Db 121 DPLNNTPTVTGASPGGLRLQLRSLEILKGVLIQRPOLCYODTILMKDIFHKNQOLA 180
QY 181 LFLIDTNRSRACHPCSPMKGSRGWSESDCQSLRTVACAGGACARCKGPLPTDCCHQC 240
Db 181 LFLIDTNRSRACHPCSPMKGSRGWSESDCQSLRTVACAGGACARCKGPLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFHFNHSGICELHCPALVTYNTDFESHPNPEGRTYFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHFNHSGICELHCPALVTYNTDFESHPNPEGRTYFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPPLHNOEYTAEDGTORCEKSKPCAR 340
Db 301 YNYLSTDVGSCTLVCPPLHNOEYTAEDGTORCEKSKPCAR 340

RESULT 4

US-09-854-356-6
; Sequence 6, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 6
; LENGTH: 919
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: fusion protein
; OTHER INFORMATION: of ECD and PD of human HER-2/neu

US-09-854-356-6

Query Match 81.1%; Score 340; DB 9; Length 919;

Best Local Similarity 100.0%; Pred. No. 1.3e-302; Indels 0; Gaps 0;

Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 METALACRMGLLALLPPGAASSTOVCCTGDMKRLRLPASPEHLMRLHLYOGCOVVOGNTL 60
DB 1 METALACRMGLLALLPPGAASSTOVCCTGDMKRLRLPASPEHLMRLHLYOGCOVVOGNTL 60

QY 61 ELTYLPFNASLSFLQDIOEVGVYLLAHNOVROVPLQRLRIYVGTQLFEDNVALAVLDNG 120
DB 61 ELTYLPFNASLSFLQDIOEVGVYLLAHNOVROVPLQRLRIYVGTQLFEDNVALAVLDNG 120

QY 121 DPLNNTPTVGTGASPGGLRELQSLREILKGVLIQIRNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVGTGASPGGLRELQSLREILKGVLIQIRNPOLCYODTILMKDIFHKNNOLA 180

QY 181 LTLIDNRSRACHPCSPMKSGKRCWGESSEDCOSLRTVACAGCARGKPLPTDCCHQC 240
DB 181 LTLIDNRSRACHPCSPMKSGKRCWGESSEDCOSLRTVACAGCARGKPLPTDCCHQC 240

QY 241 AACCTGPKHSDCLACLFHNHSGICELHCPALVYNTDFESMPNPEGRTYFGASCVTACP 300
DB 241 AACCTGPKHSDCLACLFHNHSGICELHCPALVYNTDFESMPNPEGRTYFGASCVTACP 300

QY 301 YNLTSTDVGSCTLVCPPLHNOEYTAEDGTCRCKSPCAR 340
DB 301 YNLTSTDVGSCTLVCPPLHNOEYTAEDGTCRCKSPCAR 340

RESULT 5

US-09-769-508-2

Sequence 2, Application US/09769508

Patent No. US20020155527A1

GENERAL INFORMATION:

APPLICANT: STUART, SUSAN G.

APPLICANT: MONAHAN, JOHN J.

APPLICANT: LANGTON, BEATRICE CLAUDIA

APPLICANT: HANCOCK, MIRIAM E.C.

APPLICANT: CHAO, IORRINE A.

APPLICANT: BLUFORD, PETER

TITLE OF INVENTION: C-ERBB-2 EXTERNAL DOMAIN: GP75

FILE REFERENCE: BEBIO-111-C1

CURRENT APPLICATION NUMBER: US/09/769,508

NUMBER OF SEQ ID NOS: 2

SOFTWARE: Patentlin Ver. 2.1

SEQ ID NO 2

LENGTH: 1255

TYPE: PRT

ORGANISM: Homo sapiens

US-09-769-508-2

Query Match 81.1%; Score 340; DB 9; Length 1255;

Best Local Similarity 100.0%; Pred. No. 1.7e-302; Indels 0; Gaps 0;

Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 METALACRMGLLALLPPGAASSTOVCCTGDMKRLRLPASPEHLMRLHLYOGCOVVOGNTL 60
DB 1 METALACRMGLLALLPPGAASSTOVCCTGDMKRLRLPASPEHLMRLHLYOGCOVVOGNTL 60

QY 61 ELTYLPFNASLSFLQDIOEVGVYLLAHNOVROVPLQRLRIYVGTQLFEDNVALAVLDNG 120
DB 61 ELTYLPFNASLSFLQDIOEVGVYLLAHNOVROVPLQRLRIYVGTQLFEDNVALAVLDNG 120

QY 121 DPLNNTPTVGTGASPGGLRELQSLREILKGVLIQIRNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVGTGASPGGLRELQSLREILKGVLIQIRNPOLCYODTILMKDIFHKNNOLA 180

QY 181 LTLIDNRSRACHPCSPMKSGKRCWGESSEDCOSLRTVACAGCARGKPLPTDCCHQC 240
DB 181 LTLIDNRSRACHPCSPMKSGKRCWGESSEDCOSLRTVACAGCARGKPLPTDCCHQC 240

QY 241 AACCTGPKHSDCLACLFHNHSGICELHCPALVYNTDFESMPNPEGRTYFGASCVTACP 300
DB 241 AACCTGPKHSDCLACLFHNHSGICELHCPALVYNTDFESMPNPEGRTYFGASCVTACP 300

QY 301 YNLTSTDVGSCTLVCPPLHNOEYTAEDGTCRCKSPCAR 340
DB 301 YNLTSTDVGSCTLVCPPLHNOEYTAEDGTCRCKSPCAR 340

RESULT 6

US-09-854-356-1

Sequence 1, Application US/09854356

Patent No. US20020177567A1

GENERAL INFORMATION:

APPLICANT: Cheever, Martin A.

APPLICANT: Gheysen, Dirk

APPLICANT: Corixa Corporation

APPLICANT: SmithKline Beecham Biologicals S. A.

TITLE OF INVENTION: HER-2/neu Fusion Proteins

FILE REFERENCE: 014058-009810PC

CURRENT APPLICATION NUMBER: US/09/854,356

CURRENT FILING DATE: 2001-05-09

PRIOR APPLICATION NUMBER: US 09/493,480

PRIOR FILING DATE: 2000-01-28

PRIOR APPLICATION NUMBER: US 60/117,976

PRIOR FILING DATE: 1999-01-29

NUMBER OF SEQ ID NOS: 26

SOFTWARE: Patentlin Ver. 2.1

SEQ ID NO 1

LENGTH: 1255

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: human HER-2/neu protein

NAME/KEY: DOMAIN

LOCATION: (1)..(653)

OTHER INFORMATION: extracellular domain (ECD)

NAME/KEY: DOMAIN

LOCATION: (676)..(1255)

OTHER INFORMATION: Intracellular domain (ICD)

NAME/KEY: DOMAIN

LOCATION: (990)..(1255)

OTHER INFORMATION: phosphorylation domain (PD)

NAME/KEY: DOMAIN

LOCATION: (990)..(1048)

OTHER INFORMATION: fragment of the phosphorylation domain, preferred

OTHER INFORMATION: portion (delta PD)

US-09-854-356-1

Query Match 81.1%; Score 340; DB 9; Length 1255;

Best Local Similarity 100.0%; Pred. No. 1.7e-302; Indels 0; Gaps 0;

Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 METALACRMGLLALLPPGAASSTOVCCTGDMKRLRLPASPEHLMRLHLYOGCOVVOGNTL 60
DB 1 METALACRMGLLALLPPGAASSTOVCCTGDMKRLRLPASPEHLMRLHLYOGCOVVOGNTL 60

QY 61 ELTYLPFNASLSFLQDIOEVGVYLLAHNOVROVPLQRLRIYVGTQLFEDNVALAVLDNG 120
DB 61 ELTYLPFNASLSFLQDIOEVGVYLLAHNOVROVPLQRLRIYVGTQLFEDNVALAVLDNG 120

QY 121 DPLNNTPTVGTGASPGGLRELQSLREILKGVLIQIRNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVGTGASPGGLRELQSLREILKGVLIQIRNPOLCYODTILMKDIFHKNNOLA 180

QY 181 LTLIDNRSRACHPCSPMKSGKRCWGESSEDCOSLRTVACAGCARGKPLPTDCCHQC 240
DB 181 LTLIDNRSRACHPCSPMKSGKRCWGESSEDCOSLRTVACAGCARGKPLPTDCCHQC 240

QY 241 AACCTGPKHSDCLACLFHNHSGICELHCPALVYNTDFESMPNPEGRTYFGASCVTACP 300
DB 241 AACCTGPKHSDCLACLFHNHSGICELHCPALVYNTDFESMPNPEGRTYFGASCVTACP 300

QY 301 NYLSTFVSGSCLVCPPLHNGEYTAEDGIGRCECKSKPAP 340
| | | | | | | | | | | | | | | | | | | | | | | |
Db 301 NYLSTFVSGSCLVCPPLHNGEYTAEDGIGRCECKSKPAP 340

RESULT 7
US-09-930-125-2
; Sequence 2, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:

APPLICANT: Hand-Zimmerman, Susan
APPLICANT: Cheever, Martin A.
APPLICANT: Foy, Teresa M.
APPLICANT: Lodes, Michael J.
APPLICANT: Kalos, Michael D.
APPLICANT: McNeill, Patricia D.
APPLICANT: [REDACTED]

```

: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
:
: TITLE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES
:
: FILE REFERENCE: 21021.544
:
: CURRENT APPLICATION NUMBER: US/09/930,125
:
: CURRENT FILING DATE: 2001-08-14
:
: NUMBER OF SEQ. ID NOS.: 25

```

```

: SOFTWARE: FastSeq for Windows Version 3.0
:
: SEQ ID NO 2
:
: LENGTH: 1255
:
: TYPE: prt
:
: ORGANISM: Homo sapien
:
: US-09-930-125-2

```

Query Match	81.1%	Score	340	DB	9	Length	1255
Best Local Similarity	100.0%	Pred. No.	1.7e-302				
Matches	340	Conservative	0	Mismatches	0	Indels	0
						Gaps	0

[illegible]

QY 61 ELTYLPTNASTSLFDIQEVGCVLIANHNVQRPPLQRLRIVRGQLFEDNALAVLDNG 120

Db 61 ELTYLPTNASTSLFDIQEVGCVLIANHNVQRPPLQRLRIVRGQLFEDNALAVLDNG 120

QY 121 DPLNNTTPVIGASPGSLREIQRLSLETILKGVLIQNRNQLCYDTILMKDIFHKNNQLA 180C

Db 121 DPLNNTTPVIGASPGSLREIQRLSLETILKGVLIQNRNQLCYDTILMKDIFHKNNQLA 180C

0y 181 LTLLDNRSRACHPCSPMCKSGRCWGESSEDDCSLRTVACGGCARCKGPIPTDCCHQC 240
| | | | |
Db 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDCSLRTVACGGCARCKGPPTDCCHQC 240

Oy 241 AAGCTGPKHSDDLACLHFNHSGICELHCPLALVYNTDTFESMPNPEGRTTGCASCVTACP 3000
 |||||
Dd 241 AAGCTGPKHSDDLACLHFNHSGICELHCPLALVYNTDTFESMPNPEGRTTGCASCVTACP 3000

Dy 301 YNLTSDVGSCTLVCP LHNQEVTAEDGTQRCEKCSKPCAR 340
|||
Dd 301 YNLTSDVGSCTLVCP LHNQEVTAEDGTQRCEKCSKPCAR 340

RESULT 8
US-09-441-411-6
Source: https://www.industrydocuments.ucsf.edu/docs/00441111

Publication No. US20030008342A1
GENERAL INFORMATION:
APPLICANT: Scholler, Nathalie B.
ADDRESSEE: Dais, Yusei I.

```

; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 720023 400

```

```

; CURRENT APPLICATION NUMBER: US/09/441,411
; CURRENT FILING DATE: 1999-11-16

```

```

: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: FastSeq for Windows Version 4.0
:
: SEQ ID NO: 6
: LENGTH: 1255
: TYPE: prt
: ORGANISM: Homo sapiens
:
: JS-09-441-411-6

```

Query Match	81.1%;	Score 340;	DB 9;	Length 1255;
Best Local Similarity	100.0%;	Pred. No. 1.7e-302;		
Matches 340;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1 MELALCRWGLLALLPPGAASVCTGTDMKRLRASPETHDMLRHLYQGCVVVGNL 60
Db 1 MELALCRWGILLALLPPGAASVCTGTDMKRLRASPETHDMLRHLDYGGCVVVGNL 600

QY 61 ELTYLEPTNASLSFLQDIOEVGGVYLIAHNQVRQPRLRIYRGTLFEDNAYALAVLDNG 12
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 61 ELTYLPTNASLSFLQDIOEVGGVYLIAHNQVRQPRLRIYRGTLFEDNAYALAVLDNG 12

QY 121 DPLNNTTPTVGASPGGLREQLRSLTEILKGGVLIQRNPQCYQDTILMKDIFHRNNQLA 18
 |||||
 DB 121 DPLNNTTPTVGASPGGLREQLRSLTEILKGGVLIQRNPQCYQDTILMKDIFHRNNQLA 18

Qy 181 LTLIDTNRSRACHPCSPMKGSRWGESSEDCQSLRTVCAGGACARCKPLPTDCHEQC 24
 |||||
 Db 181 LTLIDTNRSRACHPCSPMKGSRWGESSEDCQSLRTVCAGGACARCKPLPTDCHEQC 24

QY 241 AAGCGPKHSDCLACLNHHSGIELHCPLAVTYNTDTFESMNPPEGRTFGASCVTACP 30
|||
Db 241 AAGCTGPKHSDCLACLNHHSGIELHCPLAVTYNTDTFESMNPPEGRTFGASCVTACP 30

QY 301 YNLTSDVGSCTIVCPPLNQEVTAEDGIQRCECKSKPCAR 340
| | | | |
Db 301 YNLTSDVGSCTIVCPPLNQEVTAEDGIQRCECKSKPCAR 340

RESULT 9
US-09-811-123-9
Sequence 9 Application US/09811123

Sequence 9, APPLICATION 05/09611123
Patent No. US20020001387A1
GENERAL INFORMATION:
APPLICANT: Sharon Erickson
APPLICANT: Mark Szymowski
APPLICANT: Ralph Schwall
TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ERBB
TITLE OF INVENTION: ANTIHER-1/METASTASINOID COMPOUNDS

FILE REFERENCE: GENEHT 07/342
CURRENT APPLICATION NUMBER: 05/09/811,122
CURRENT FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/238,327
PRIOR FILING DATE: 2000-10-05
PRIOR APPLICATION NUMBER: 09/602,530
PRIOR FILING DATE: 2000-06-23

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1255
END
```

ORGANISM: Homo sapiens
US-09-811-123-9

Best Local Similarity	100.0%	Pred. No. 1.7e-302
Matches	340	Conservative 0; Mismatches 0; Indels 0; Gaps 0

Db 1 MELALCRWGLLALLPPGAASYCTGTGDMKRLRLEASPETHDMLRHLYGGCQVQGNL 600

Db 61 ELTYLPNTASLSFLDIOEVGCVLIAHNOVROPFLQRRLRYRGTLQFEDNTALAVLDNG 122

US-09-821-883-4

Query Match 69.7%; Score 292; DB 10; Length 697;

Best Local Similarity 100.0%; Pred. No. 7.6e-259; Mismatches 0; Indels 0; Gaps 0;

Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 19 GAASTOVCTGTDKRLRPASPEHLDMLRHLVQGCVOVQGNLELYLPTNASLSFLQDIO 78
    |||||||
Db 38 GAASTOVCTGTDKRLRPASPEHLDMLRHLVQGCVOVQGNLELYLPTNASLSFLQDIO 97
    |||||||
QY 79 EVQGYVLAHNGVROVPLQRLRIVRGTOLEEDNYALAVLDNGDPLNNTTPYTGASPGGLR 138
    |||||||
Db 98 EVQGYVLAHNGVROVPLQRLRIVRGTOLEEDNYALAVLDNGDPLNNTTPYTGASPGGLR 157
    |||||||
QY 139 ELQLRSLEILKGVLIQRNPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPM 198
    |||||||
Db 158 ELQLRSLEILKGVLIQRNPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPM 217
    |||||||
QY 199 CKGSRGWGSESSDCSLRTYVCAGGCARCKGPLPTDCHEQCAAGCTGPKHSDCLACLHF 258
    |||||||
Db 218 CKGSRGWGSESSDCSLRTYVCAGGCARCKGPLPTDCHEQCAAGCTGPKHSDCLACLHF 277
    |||||||
QY 259 NMSGICELHCPALVYNTDFESMPNPEGRTYTFGASCVTACPYNYLSTDVGS 310
    |||||||
Db 278 NMSGICELHCPALVYNTDFESMPNPEGRTYTFGASCVTACPYNYLSTDVGS 329
    |||||||
```

Search completed: March 4, 2003, 12:54:28
Job time : 20.6687 secs

